RICE UNIVERSITY

Ligand Diffusion Pathways and Mechanisms for Regulating O_2 Affinity in Two Model Invertebrate Globins: The E7 Gate and Apolar Tunnel

by

Mallory D. Salter

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

John S. Olson, Ralph and Dorothy Looney Professor
Biochemistry and Cell Biology, Director

Edward P. Nikonowicz, Associate Professor
Biochemistry and Cell Biology

Michael Kohn, Assistant Professor
Ecology and Evolutionary Biology

Jonathan Silberg, Assistant Professor
Biochemistry and Cell Biology

HOUSTON, TEXAS
APRIL 2011
ABSTRACT

Ligand Diffusion Pathways and Mechanisms for Regulating O_2 Affinity in Two Model Invertebrate Globins: The E7 Gate and Apolar Tunnel

By

Mallory D. Salter

The major pathway for O_2 binding to mammalian myoglobins (Mbs) and hemoglobins (Hbs) involves transient outward movements of the distal histidine (HisE7), which allows ligand migration into the distal portion of the heme pocket. This E7 gate pathway appears dominant in vertebrate Hbs and Mbs. However, a number of invertebrate globins, including the dimeric hemoglobin from the blood clam *Scapharca inaequivalvis* (ScHbI), have an inverted quaternary structure in which an EF:FE dimer interface appears to block the HisE7 gate. Another set of globins, including the mini-hemoglobin from the Nemertean sea worm *Cerebratulus lacteus* (CerHb), are missing the N-terminal A-helix, which results in an internal tunnel between the E- and H-helices. This apolar channel has been suggested to represent an alternative to the HisE7 gate pathway. To determine the roles of the E7 gate and alternative pathways, we have systematically examined the effects of mutations at the E7 position in ScHbI and CerHb and at 21 other positions along the polar channel in CerHb.

As was observed for SwMb and HbA, there is a progressive decrease in the bimolecular rate constants for O_2 binding to ScHbI as the size of the amino acid at position E7 is increased from Ala to Trp. This pattern is unaffected when ScHbI is completely converted to the R- or high affinity quaternary state by the F97Y mutation or when the dimer interface is completely disrupted by the K30D mutation. In contrast, E7 mutations have little affect on the rates of ligand entry and escape in CerHb. Instead, ligands diffuse through the apolar channel between the E- and H-helices as judged by

decreases in both overall association and dissociation rate constants and increases in the extent of geminate recombination when the channel is blocked by small to large amino acid mutations. In SwMb, these trends are only observed when the small to large mutations are constructed at or near the E7 gate or directly in the distal pocket where ligands are captured. Thus, it is clear that globins have evolved more than one pathway for rapid O_2 uptake and release.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. John Olson. He not only has a fervent passion for researching heme proteins, but also is most dedicated to teaching his students, undergraduates and graduates. He gives a tremendous amount of his time and attention to each of his students, and we are ever thankful.

I would also like to thank Eileen Singleton for constructing, expressing, and purifying the over 200 mutant proteins described in my thesis. With her diligence, I was able to sum up work started in 1988 on sperm whale myoglobin and in 1998 on *Cerebratulus lacteus*.

Other graduate students in the Olson Lab that I interacted with over the years include Ivan Birukou, Todd Mollan, David Culbertson, George Blouin, Angela Hvitved, and Philip Graves. The camaraderie between students in the Olson Lab made for a healthy and entertaining environment. George Blouin taught me the techniques of laser flash photolysis as well as FTIR, and Todd Mollan assisted me with cloning techniques. I am grateful for their efforts and in taking time out of their own schedules to work with me. As a graduate student in the Olson Lab, I formed lasting relationships, enduring tough times and celebrating in life's accomplishments. These impressions will remain on my heart for a lifetime.

I would also like to thank my sister, Meghan Roberts, for not only taking care of my son for many months, but in being a close friend and a voice of uplifting thought. I would also like to profess my gratitude to my three moms, Barbara Dorand, Pat Salter, and Angela Salter. They not only supported my husband and I in our first years of marriage as I entered into graduate school, but they also provided many days of child care

for both my sons. Without their flexibility and willingness to help, graduate school may have taken a little longer!

I would also like to thank my friends, husband, and children. My friends outside of the Rice community have given me a glimpse that life goes on outside of school. As a traditional student, passing straight from college to graduate school, there were many times that I felt I had been in school 'forever'. It seemed at times, a better choice would have been to just stop and get a job. My friends encouraged me, and graduate school is where I remained.

My husband also encouraged me to no avail. We married in February of 2006, and moved straight to Houston in May to begin graduate school. Until now, our marriage has included 5 years of graduate school, 6 moves across Houston, 2 houses, 2 children, and much more. I am extremely humbled by his love and adoration for our children and me, and I am extremely grateful for his dedication to work and taking care of his family.

My children, Robert West (November 16, 2010) and James Reese (March 18, 2009) are both wonderful people and tremendous blessing in our lives. They have each changed my life in many ways, and I am very thankful for their unique personalities and the joy they have brought to our lives.

Lastly, I would like to acknowledge my faith in God. He has not only given me the strength to endure, but also the talents that have flourished throughout my education. I am forever grateful, and humbled to be one of his children (Isaiah 40: 31).

Table of Contents

List of Figures

List of Tables

List of Abbreviations, Conventions, Symbols

ALA: δ–aminolevulinic acid

AscHb: *Ascaris suum* domain I hemoglobin

Atm: Atmosphere

CerHb: *Cerebratulus lacteus* mini-hemoglobin

CHES: 2-(N-Cyclohexylamino)ethane sulfonic acid

CM: Carboxymethyl

CO: Carbon monoxide

Cygb: Cytoglobin

DEAE: Diethylaminoethyl

DNase I: Deoxyribonuclease I

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

FAD: Flavin adenine dinucleotide

$FeSO_4$: Iron (II) sulfate

F_{gem}: Fraction of geminate recombination

FHb: Flavohemoglobin

FTIR: Fourier transform infrared spectroscopy

FTIR-TDS: Fourier transform infrared spectroscopy- temperature derivative spectroscopy

GCS: Globin coupled sensor

Hb: Hemoglobin

HbA: Adult human hemoglobin A

HemAT: Heme-based aerotactic transducers

Hemin: Fe(III)-protoporphyrin IX

HEPES: 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

IPTG: Isopropyl-1-thio-b-D-galactopyranoside

k'_X: Ligand association rate constant

K_2HPO_4: Dipotassium phosphate

k_{gem}: Geminate recombination rate constant

KH_2PO_4: Monopotassium phosphate

KPi: Potassium phosphate

K_X: Ligand affinity

k_X: Ligand dissociation rate constant

LB: Luria broth

Lba: Leghemoglobin

Mb: Myoglobin

MD: Molecular dynamics

$MgCl_2$: Magnesium chloride

NaCl: Sodium chloride

NaPi: Sodium phosphate

$\nu_{C\text{-}O}$: Wavenumber, FTIR C-O stretching frequency

Ng: Neuroglobin

NO: Nitric oxide

NsHb: Nonsymbiotic hemoglobin

O_2: Dioxygen

PMFS: Phenylmethanesulfonylfluoride

RNase A: Ribonuclease A

ScHbI: *Scapharca inaequivalvis* hemoglobin domain I

SDG: Single domain globin

SDS-PAGE: Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SwMb: Sperm whale myoglobin

TB: Terrific broth

TrHb: Truncated hemoglobin

TrHbN: Truncated hemoglobin from *Mycobacterium tuberculosis*

Tris-HCl: Tris(hydroxymethyl)aminomethane

WT: Wild type

Xe: Xenon

YAG: Yttrium aluminum garnet

Equations

Chapter 1: Introduction and Background

1.1 General introduction - Mammalian myoglobin (Mb) and hemoglobin (Hb) are well-known O_2 storage and delivery heme proteins (*1-3*) and have been the topic of much research since Perutz and Kendrew first reported their three-dimensional structures over 45 years ago (*4, 5*). In the last twenty years, globins have been discovered in all kingdoms of life, expanding this family of proteins from the previously known vertebrate α- and β- globins of human Hb, mammalian Mbs, invertebrate Hbs, and the symbiotic leghemoglobins found in leguminous plants (Lbas) (*6*). For instance, nonsymbiotic Hbs (NsHbs) have been found in all plants (*7-9*) and chimeric flavohemoglobins (FHbs) have been found in bacteria and fungi. These FHbs are examples of newly defined multidomain globins within the Hb superfamily composed of an N-terminal globin domain linked to an FAD reductase domain and a C-terminal domain resembling simple bacterial Hbs (*10*). Two additional globins discovered within the animal kingdom are cytoglobin (Cygb) and neuroglobin (Ng) that appear to be ubiquitous among all vertebrates even though their functions are still under debate (*11*).

Most of these single and multidomain globins belong to the 3/3 α-helical fold family (Fig 1.1A). This structure consists of seven to eight α-helices (designated A through H) in which two sets of three α-helices (BEF/AGH) wrap around a heme prosthetic group (Fig 1.1B). A newer family of globins referred to as truncated hemoglobins (TrHbs) have a 2/2 α-helical fold (Fig. 1.1C) in which two sets of two α-helices (BE/GH) wrap around a heme prosthetic group (*12*). TrHbs are typically 20-40 residues shorter than most mammalian Hbs (*12*), and have been found in green algae (*13*), plants (*14*), protozoa (*15*) and bacteria (*16*).

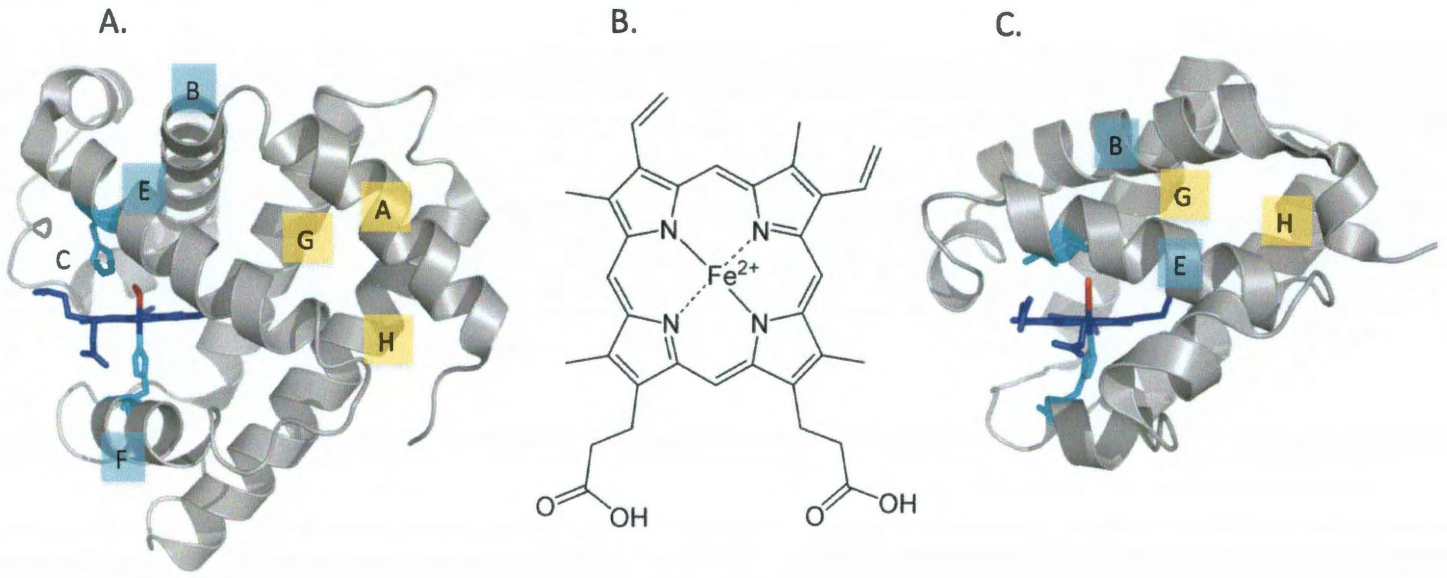


Figure 1.1 Structural variation among the globin superfamily. *A*, Model of a 3/3 α-helical fold illustrated by sperm whale myoglobin (PDB ID: 2mgm). *B*, Fe (II) protoporphyrin IX (heme b). *C*, Model of a 2/2 α-helical fold illustrated by *Bacillus subtilis* hemoglobin (PDB ID: 1ux8). The globin backbones are designated by gray cartoon, the E7 and F8 positions highlighted by cyan sticks, the heme groups by blue sticks, and bound oxygen and cyanide by red sticks in *A* and *C*, respectively. The protein structures were generated in MacPyMol and heme b in ChemDraw.

The discovery of globins in all kingdoms of life and animal subphyla led to the recognition of a variety of globin primary and quaternary structures. For example, just among invertebrate globins, the quaternary state can range from single domain/ monomeric globins, exemplified by the mini-hemoglobin from *Cerebratulus lacteus* (109 amino acids), to > 3 million Dalton (Da) aggregates of dodecameric globins, exemplified by *Lumbricus terrestris* hemoglobin (Fig. 1.2). In general, monomeric Hbs include the single domain globins (SDGs) of approximately 17 kDa molecular weight with either the traditional 3/3 α-helical fold or the 2/2 α-helical fold.

The specific structural features observed in each globin family finely tune the physiological role of the individual globin. The main functions of most animal muscle and blood cell hemoglobins involve O_2 storage and transport. Recently, a number of researchers have shown that NO dioxygenation (the conversion of nitric oxide to nitrate by MbO_2 or HbO_2; Fig. 1.3) is an important secondary physiological function in both these globins (*17-21*). Since nitric oxide can irreversibly disrupt the Fe/S center in aconitase (*22*) and bind to cytochrome c oxidase with high affinity (*18*), NO is highly toxic to both neuronal and striated muscle cells because it inhibits mitochondrial respiration. Once the NO radical is captured in the distal pocket[1] of globins, it reacts immediately with bound dioxygen, which resembles superoxide having a partial negative charge. As a result, the rate of NO dioxygenation in myocytes and red cells is directly related to the rate at which NO enters the reactive Fe center of the heme prosthetic group (Fig. 1.3).

[1] The distal pocket refers to the cavity surrounding the bound ligand. The distal pocket is opposite the proximal pocket which contains the His(F8) residue that forms the only covalent bond to the heme prosthetic group.

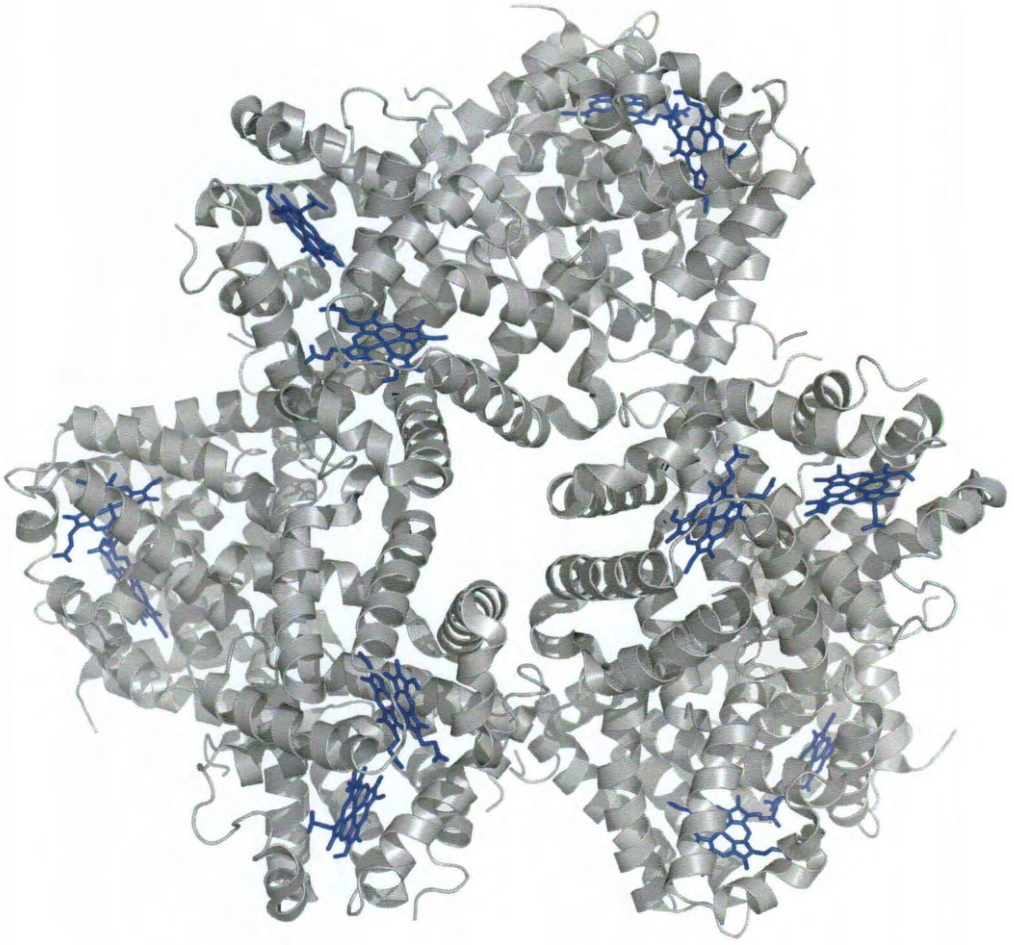


Figure 1.2. Structural representation of the dodecameric hemoglobin unit from *Lumbricus terrestris* (PDB ID: 1x9f). The globin backbone is represented by gray cartoon and each heme group is designated by blue sticks. The protein structure was generated in MacPyMol.

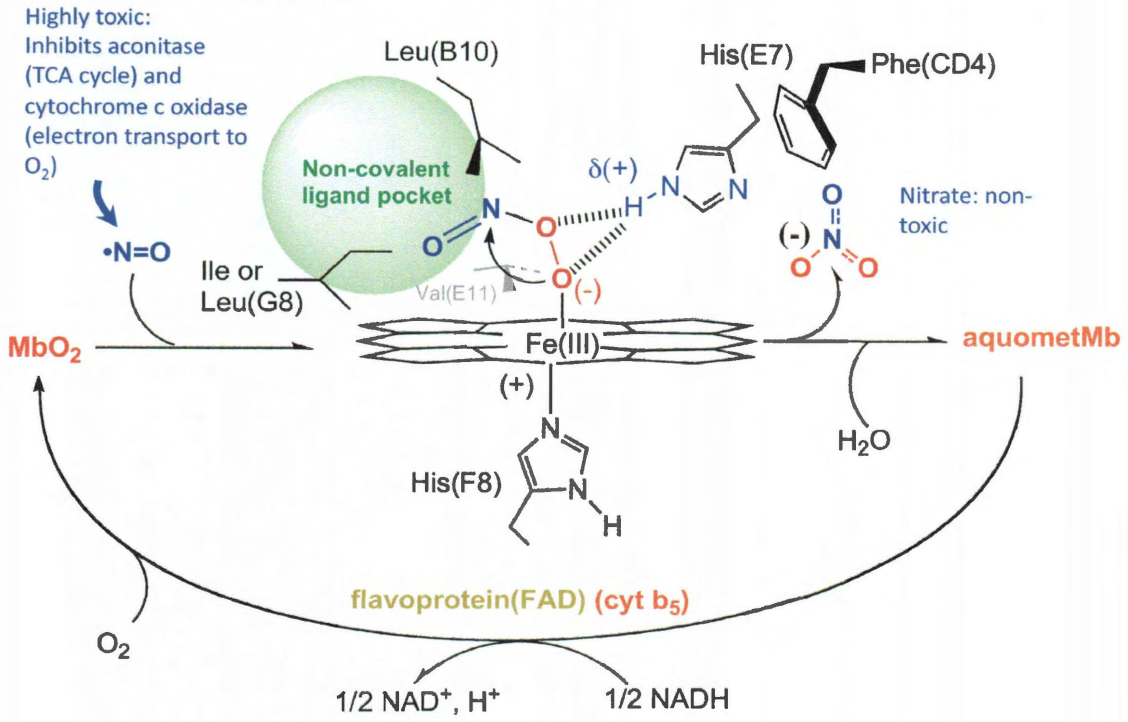


Figure 1.3. Nitric oxide dioxygenation by oxymyoglobin. NO dioxygenation to oxymyoglobin occurs by temporary disruption of the hydrogen bond between the distal histidine (HisE7) and bound oxygen followed by rapid reaction with coordinated oxygen. Nitric oxide, a highly toxic molecule, is thus converted to nitrate, a non-toxic product. This figure was taken from the proposed scheme based on work from (*20, 21, 23-25*).

Additional functions of Hbs include: O_2 sensing in globin coupled sensors (GCS) such as the heme-based aerotactic transducers (HemATs) found in Archaea and Bacteria (*26*); O_2 scavenging or sequestering in anaerobic organisms including parasitic roundworms and bacteria having anaerobic growth cycles; and dehaloperoxidase activity in certain sea worms ((*27, 28*); see review by (*29*)).

1.2 Ligand binding mechanism in myoglobin: the baseball glove model - Following determinations of the initial three-dimensional structures of sperm whale Mb (SwMb) and horse heart Hb, Perutz and Matthews proposed that ligands enter globins through a short channel, which is gated by the distal E7[2] histidine and connects the reactive Fe center of the heme prosthetic group to the solvent phase surrounding the protein (*30*). This short channel is now referred to as the E7 gate pathway. In the view of most experimentalists, ligand binding to Mb is well understood ((*1, 31-37*) and references therein) and consists of four major steps (Fig. 1.4). (1) Weakly bound water is displaced to create a vacant distal pocket above the heme Fe atom. (2) Ligands migrate into the protein through a short channel that is created when the distal histidine (His64) transiently rotates upward and then are captured in the interior pocket of the active site. This non-covalent intermediate is often called the B state because it can also be generated by photolysis of the equilibrium bound or A state. (3) Covalent bond formation between the internal ligand and the Fe atom of the heme group then competes with ligand escape back out through the HisE7 gate. (4) The bound ligand can be further stabilized by electrostatic interactions with surrounding polar amino acid side chains (Fig. 1.4).

---

[2] Amino acids have been identified with their topological site numbers and/or helical position as defined in the conventional globin fold, *i.e.* His64(E7) refers to the histidine at residue 64, which is located at the 7th position of the E-helix.

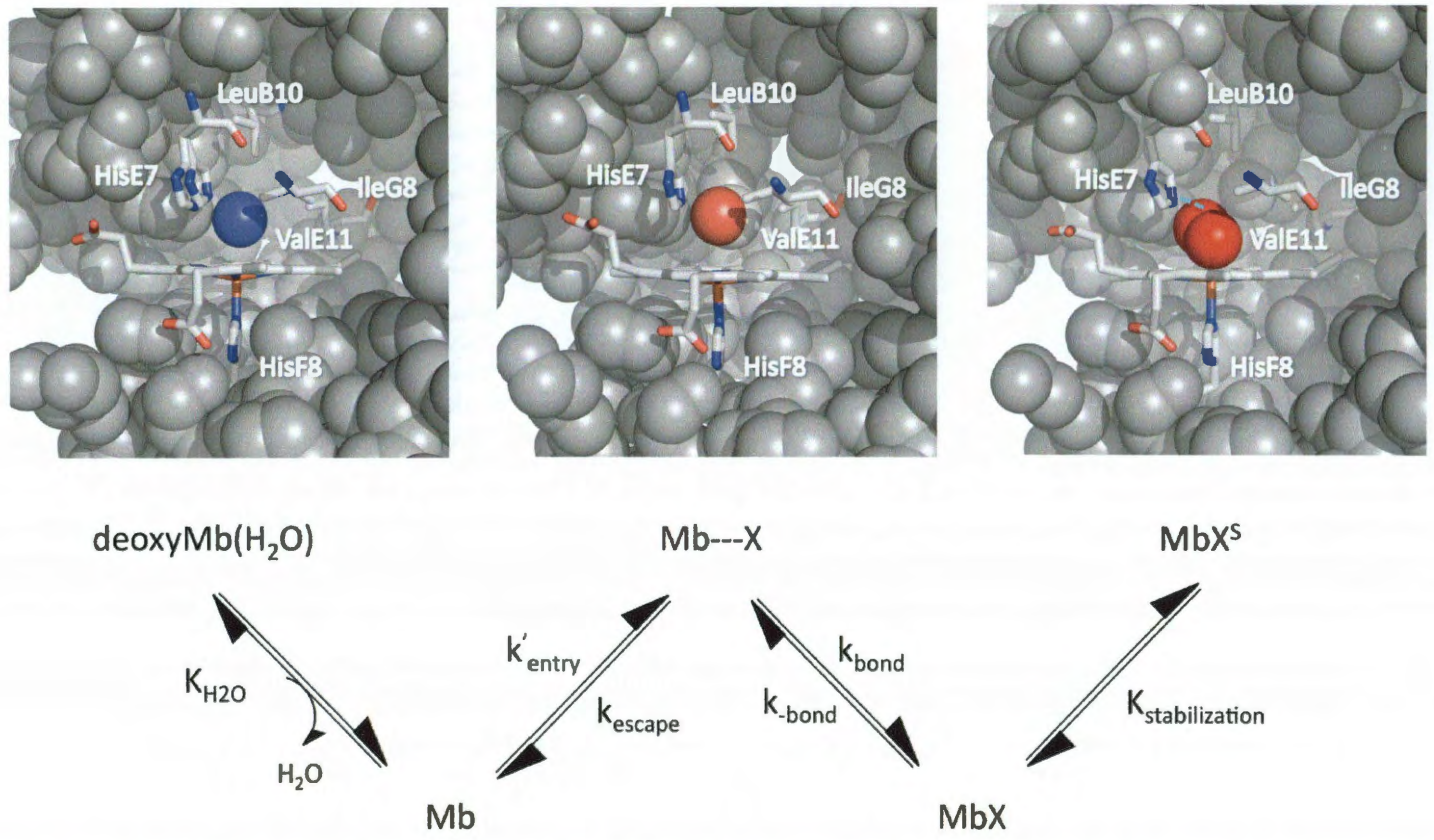


Figure 1.4. Ligand binding mechanism for sperm whale myoglobin as a model system for mammalian Mbs and Hbs. In deoxyMb (PDB ID: 1a6n), a noncovalently bound water (blue sphere) is located within the ligand-binding pocket and is displaced in the first step of binding. Residues E7, B10, E11 and G8 surround the distal pocket and HisF8 designates the proximal histidine. In the second step, the ligand (CO, in this example, designated by red and white spheres) enters the "empty" binding pocket (PDB ID: 1abs). In the third step, a covalent bond is formed between the ligand (O_2, in this example, designated by red spheres) and Fe^{2+} (PDB ID: 2mgm). In the last step, a hydrogen bond (cyan dashed line) forms between the HisE7 and the ligand to stabilize the bound ligand. Figure taken from (*33, 38*). The protein structures were generated in MacPyMol.

In this model, the distal pocket of the globin is compared to the pocket of a baseball glove. In order to 'catch' an incoming ligand, the thumb of the glove (HisE7) must swing open momentarily to allow entry of the ligand into the pocket (space between IleB9, LeuB10, ValE11, and IleG8). Following capture in the distal pocket, the histidine "thumb" rotates downward toward the heme trapping the ligand close to the reactive Fe center of the heme prosthetic group. Then a covalent bond is formed between the Fe and ligand, and finally a hydrogen bond is formed between the HisE7 and the bound ligand. Despite all the experimental evidence supporting ligand movement into Mb through this E7 pathway, most theoretical simulations suggest that multiple alternative routes are used more frequently (*39*). These theoretical simulations are further described in the introduction of Chapter 7.

1.3 Focus of thesis research - In order to gain a greater understanding of ligand diffusion pathways in globins from subphyla other than Vertebrata, the dimeric hemoglobin from the mollusk *Scapharca inaequivalvis* (ScHbI) was first explored. ScHbI has been studied for many years, and Royer *et al.*. (*40, 41*) have suggested that the specific quaternary structure in this globin might hinder or block the E7 gate pathway. Secondly, I wanted to explore a globin whose tertiary structure suggested an alternate pathway through a putative apolar tunnel (*42, 43*). In this case, I chose the mini-hemoglobin from a Nemertean sea worm, *Cerebratulus lacteus* (CerHb). Our goal was to test whether or not these two invertebrate hemoglobins employ the E7 gate pathway for ligand diffusion. If not, then a rational mutagenesis study would be used in order to map any novel pathways. In the end, the comparison between the ligand diffusion pathways in these two invertebrate hemoglobins and in SwMb has yielded a more general

understanding of ligand diffusion pathways for all globins and the selective pressures involved in the evolution of alternative pathways.

1.4 Structure of Scapharca inaequivalvis hemoglobin (ScHbI) - ScHbI is an invertebrate globin from the blood clam *Scapharca inaequivalvis* and has a distal pocket composition very similar to that of vertebrate hemoglobins, containing MetB10, PheCD1, HisE7 and LeuE11 (Fig. 1.5). This mollusk protein has an "inverted" dimeric quaternary structure compared to adult human hemoglobin A (HbA) and the E7 gate pathway in ScHbI appears to be blocked by the highly polar EF:FE dimer interface (*40, 41*) (Fig. 1.6A). This dimer interface is a common characteristic of multimeric invertebrate globins with "normal" HisE7 containing active sites but with inverted quaternary structures compared to vertebrate Hbs (*44, 45*). ScHbI was selected not only for this unique structural arrangement, but also because an expression system had been established (*46*), multiple crystal structures had been determined (*47, 48*), and extensive kinetic studies had already been published for both wild type (wt) ScHbI and a few mutants (*47, 49-51*).

In addition, ScHbI dimer can be locked into a completely R, high affinity state (*51*). In wt ScHbI, the F97 residue packs next to the proximal histidine hindering, the Fe from moving into the heme plane due to a more eclipsed heme Fe orientation (Fig. 1.7A). Upon ligand binding, F97 rotates towards the dimer interface disrupting a well-ordered cluster of water molecules, and the plane of the imidazole ring becomes more staggered in relation to the pyrrole nitrogen atoms, and, as a result, the Fe atom can readily move into the heme plane (Fig. 1.7B). These structural changes do not occur when Tyr is present at the 97(F4) position because the additional phenol hydroxyl O atom prevents

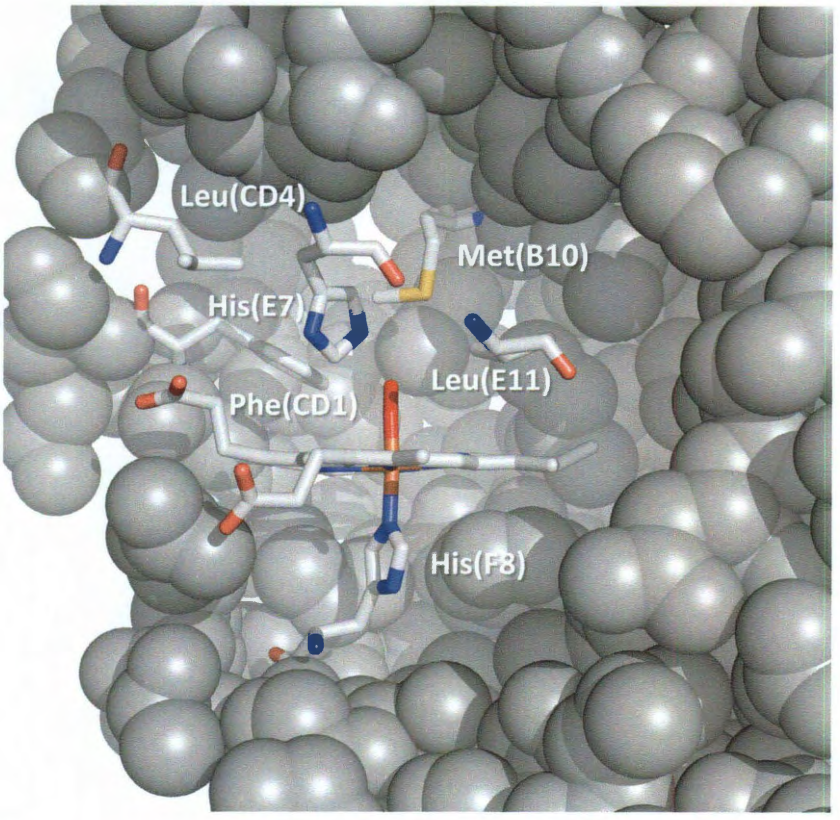


Figure 1.5. Distal pocket structure of *Scapharca inaequivalvis* oxyhemoglobin (PDB ID: 1hbI). The globin backbone is represented by gray spheres and key active site residues, MetB10, PheCD1, LeuCD4, HisE7, and LeuE11, are depicted in sticks. The perspective of this representation is across the dimer interface. One subunit has been removed and part of the protein matrix has been slabbed away to provide an unhindered view of the active site. The protein structure was generated in MacPyMol.

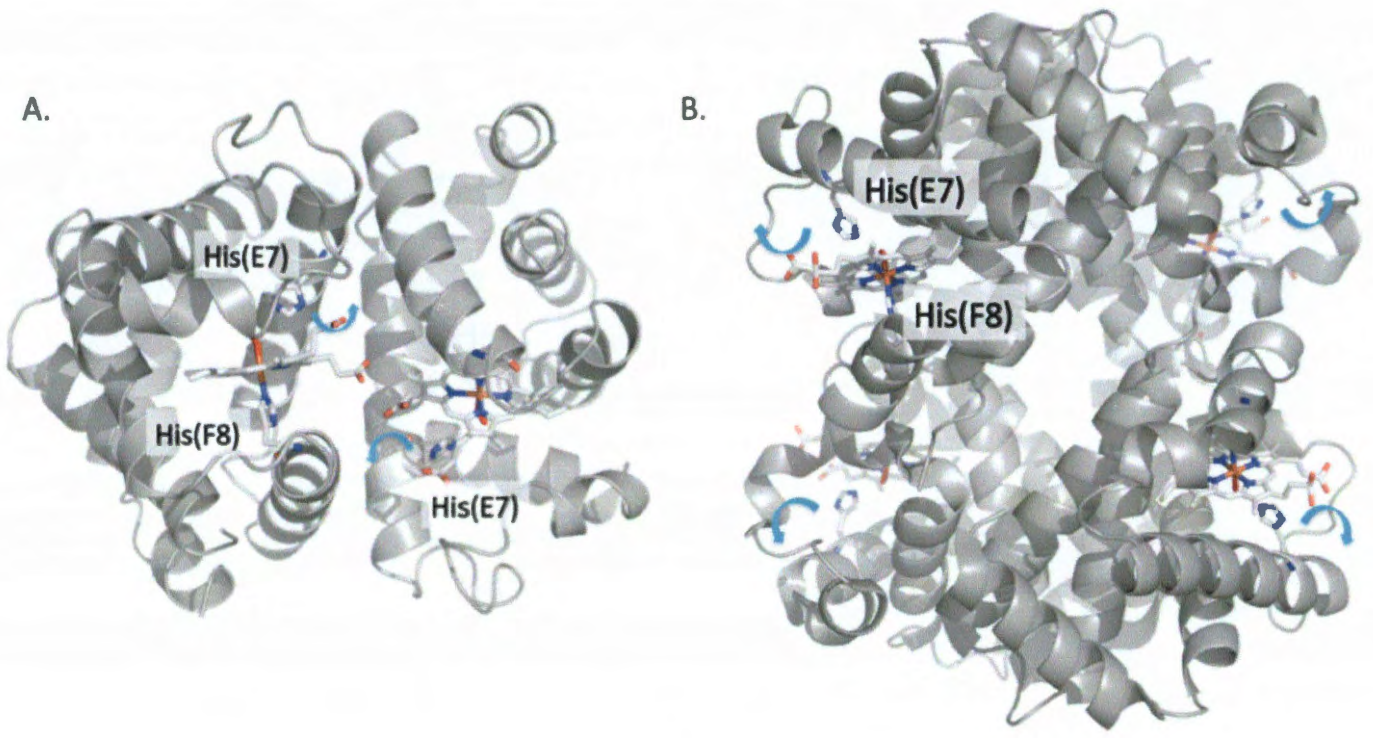


Figure 1.6. Quaternary structural representation of *Scapharca inaequivalvis* oxyhemoglobin (ScHbI) (PDB ID: 1hbI) and human hemoglobin A (HbA) (PDB ID: 2dn1). *A*, Dimeric structure of ScHbI. *B*, Tetrameric structure of HbA. In each, the globin backbone is represented as gray helices while the heme prosthetic group, HisE7, and HisF8 are depicted as sticks. In ScHbI, the HisE7 gates open towards the dimer interface as illustrated by the cyan arrows. In HbA, the HisE7 gates open towards the solvent as depicted by the cyan arrows. The protein structures were generated in MacPymol.

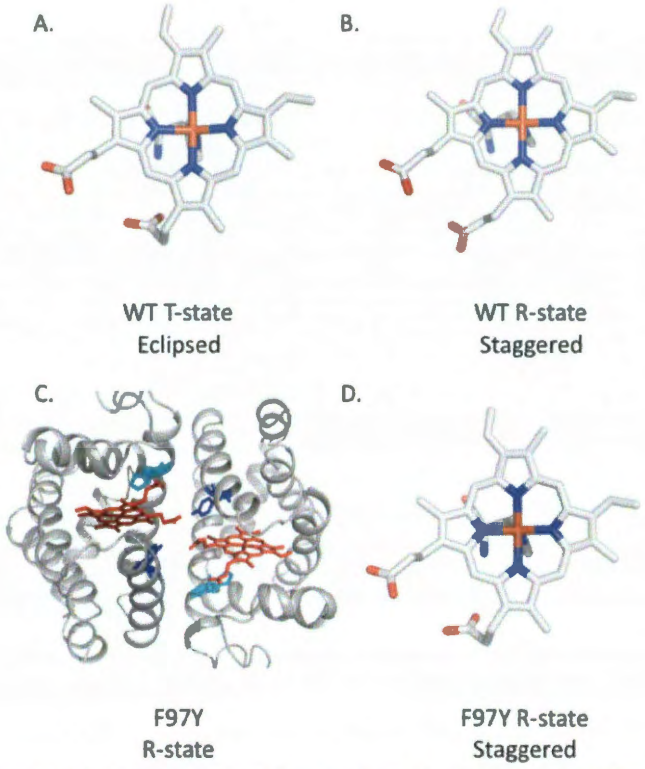


Figure. 1.7. Heme Fe orientation and proximal His geometry of T- and R-state ScHbI and F97Y ScHbI protein structure. *A*, In the unliganded or T- quaternary state, the plane of the imidazole ring is hindered from moving into the plane of the heme due to an eclipsed geometry (PDB ID: 1nwi). *B*, In the liganded or R- quaternary state, the plane of the imidazole ring is readily available to move into the plane of the heme due to a more staggered geometry (PDB ID: 1hbi). Thus, the reactivity of the Fe and the affinity for ligand increases in the R-state. *C*, Structural representation of the Phe97(F4)Tyr ScHbI mutant. The globin backbone is represented by gray cartoon, the heme group is designated by red sticks, the HisE7 residue by cyan sticks and the F97Y mutation by blue sticks. *D*, In the F97Y ScHbI mutant, the plane of the imidazole ring is readily available to move into the plane of the heme due to a constantly staggered geometry (PDB ID: 2aup). The protein structures were generated in MacPyMol.

the side chain from packing underneath the heme in the T, low affinity state, and the dimer is locked into the R state even when no ligands are bound (Fig. 1.7C). Thus, in the Tyr97(F4) mutant, the Fe is held in the more reactive, high affinity state, and, as a result, the route of entry in ScHbI can be focused on without having to analyze the complex kinetics associated with cooperative ligand binding.

1.5 Structure of Cerebratulus lacteus mini-hemoglobin (CerHb) – In contrast to animal Hbs, the pathways for ligand binding in trHbs appear to be distinct from the E7 gate found in most classic 3/3 α–helical globins (*15, 52-54*). These bacterial and plant Hbs display truncation of elements found in the classic globin fold, including almost complete loss of the A-helix, significant shortening of the CD-D helical region, and reduction of the F-helix to an extended polypeptide loop (*12, 15, 54*). The resulting tertiary structure has the appearance of a 2/2 α–helical fold (Fig. 1.1C). The abbreviated CD-E region results in the pulling of the E-helix closer to the heme distal pocket, which may preclude the use of the E7 residue as a gate for ligand entry and exit. The crystal structures of *M. tuberculosis* trHb (TrHbN and TrHbO) and other trHbs display the presence of a tunnel through the protein matrix between the E- and H- helices that is large enough for ligand diffusion to the heme distal pocket from external regions near the EF corner and the C-terminus of the protein (*12, 16, 54-57*). However, very little experimental work has been done to map these alternative pathways in trHbN (*58*) or trHbO (*59*).

A very similar protein matrix tunnel is found in the neuronal mini-Hb of the Nemertean sea worm, *Cerebratulus lacteus* (Fig. 1.8A and B), which is the smallest

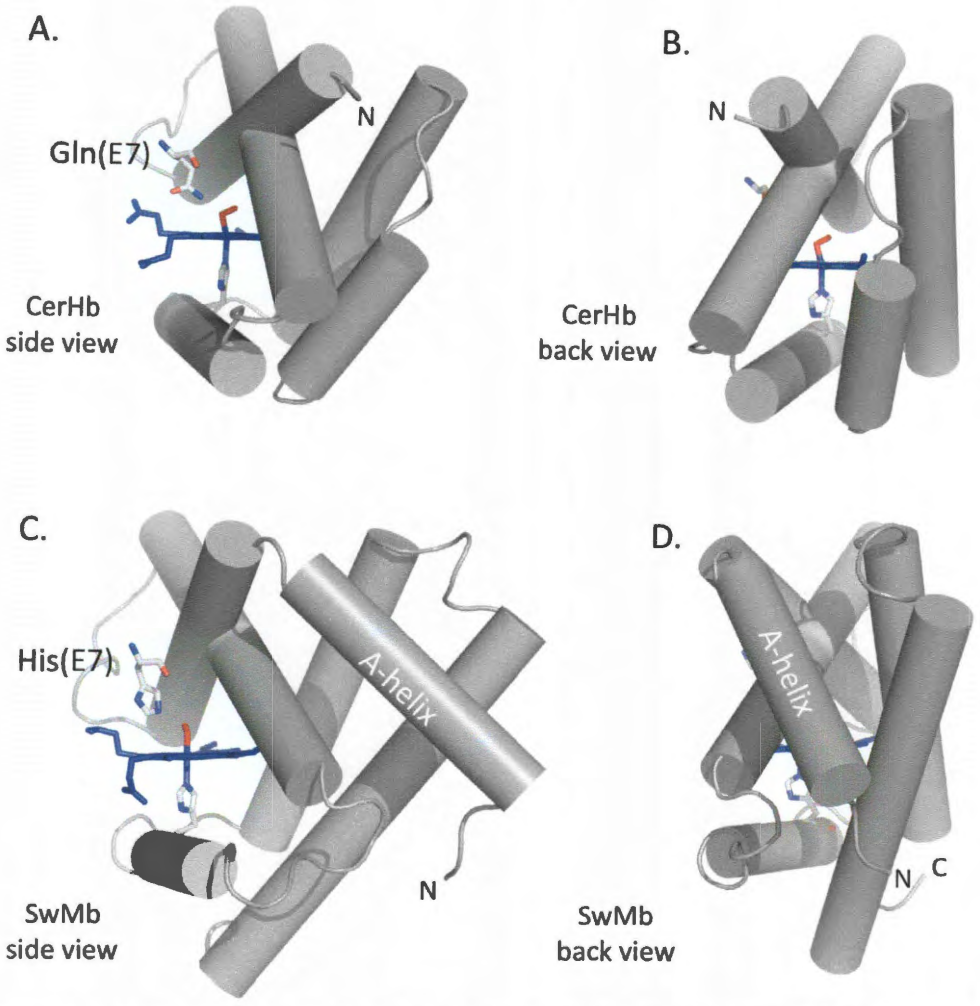


Figure 1.8. Structural representation of *Cerebratulus lacteus* mini-hemoglobin (PDB ID: 1kr7) and sperm whale myoglobin (PDB ID: 2mgm) from side and back perspectives. *A*, CerHb side view. *B*, CerHb back view. *C*, SwMb side view. *D*, SwMb back view. In each panel, the globin backbone is, represented as gray cylinders, the heme prosthetic group by blue sticks and E7 and F8 residues as sticks. The N-terminus is noted in each panel to emphasize the loss of the A-helix in CerHb. The protein structures were generated in MacPymol.

known globin with only 109 amino acids (*42, 43, 60*). Although CerHb is genetically distinct from the trHbs, it is also missing an A-helix and has marked shortening of the B- and H-helices. CerHb also has a distal pocket very similar to some truncated, microbial, and invertebrate Hbs that have TyrB10 and GlnE7 in their active sites (*42, 43*). These two polar residues normally donate strong hydrogen bonds to bound ligands that typically cause an extremely high oxygen affinity (*60, 61*). However, CerHb contains a Thr at the E11 position, and the Thr48(E11) β hydroxyl O atom pulls the TyrB10 phenol H atom away from the bound ligand. As a result, the non-bonding electrons on the TyrB10 O atom point toward the bound ligand, and this negative electrostatic field decreases the oxygen affinity of wt CerHb almost 1,000-fold to a value, K_{O2}= 1.3 μM^{-1}, similar to vertebrate myoglobins (K_{O2}= 1.1 μM^{-1}) ((*61, 62*); See Chapter 4). As a result, this CerHb has a Mb-like O_2 affinity and serves an O_2 storage function in the axons and brains of the organism. Pesce et al. (*62*) suggested that the naturally occurring ThrE11 in CerHb represents an example of convergent evolution from a high to low affinity hemoglobin to provide O_2 delivery to brain and neuronal tissue.

Despite its moderate oxygen affinity, CerHb displays very rapid rates of association (k'_{O2}= 240 $\mu M^{-1}s^{-1}$)[3] and dissociation (k_{O2}= 180 s^{-1}). These large rate constants seem surprising considering the crowded active site and multiple hydrogen bonding interactions that appear to "tie" down the amide side chain of GlnE7, which forms the E7 gate for CerHb. Bolognesi and coworkers (*42, 53, 62*) suggested that the large apolar channel created by the absence of an N-terminal A-helix and the reduction of the H-helix to 8 amino acids may be the pathway for ligand movement into the distal

[3] The term k'_{ligand} is used to describe the association rate constant of a ligand, which is defined in the subscript. Likewise, k_{ligand} refers to the dissociation rate constant.

pocket and account for the high rates of ligand binding and release. CerHb was selected for detailed study because of these unique properties, and because an efficient expression system had already been established (*43*), crystal structures had been determined (*12, 42, 60, 62*), and initial kinetic studies of wt and mutants had been published (*62*).

1.6 Thesis objectives - In this thesis work, ligand diffusion pathways were investigated in both ScHbI and CerHb. ScHbI represents a model for a 3/3 invertebrate globin with a putative E7 gate despite its inverted quaternary structure, while CerHb represents a model for a 2/2 invertebrate globin with an alternate, apolar tunnel pathway. In addition to determining if the apolar tunnel in CerHb serves as the ligand diffusion pathway, I also carried out a more detailed comparison map with SwMb by expanding the existing library of SwMb mutants to helical positions that are important in CerHb. This expansion allowed more specific comparisons between mutations in the pathways for ligand migration in both SwMb and CerHb. Our goals were to determine: (1) if the presence of His at E7 is an indicator of the E7 gate pathway in subphyla other than Vertebrata, despite variation in quaternary structure, and (2) if the TyrB10/GlnE7 motif is an indicator of a hindered E7 gate which led to the evolution of alternative pathways to maintain high rates of uptake and release.

Chapter 2: Materials and Methods

2.1 Materials - All cDNA primers were purchased from Sigma-Aldrich (St. Louis, MO); Quikchange mutagenesis kits were purchased from Stratagene (La Jolla, CA); the QIAprep Spin Miniprep kits were purchased from QIAGEN (Valencia, CA); all XL1-Blue, BL21(DE3), and BL21(DE3)pLysS competent cells were purchased from Stratagene (La Jolla, CA). W3110*lac*Iq L8 (*63*) competent cells were a gift from the Royer group (U Mass Med Center). Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was purchased from Gold BioTechnology, Inc. (St. Louis, MO); δ–aminolevulinic acid (δ–ALA) was purchased from Sigma-Aldrich (St. Louis, MO); hemin was purchased from either Porphyrin Products (Logan, UT) or Sigma-Aldrich (St. Louis, MO); and cylinders of O_2, CO, and NO were purchased from Matheson Tri-gas (Parsippany, NJ) in ultra high purity form.

2.2 ScHbI sample preparation – Wild type ScHbI was originally isolated as described in Chiancone *et al.* (*64*) and the plasmids were expressed in *E. coli* W3110*lac*Iq L8 using the scheme to express recombinant wt ScHbI as described by Summerford *et al.* (*46*). I am grateful for the wt, F97Y, and K30D plasmids, which I received from the Royer group upon initiation of this study (*46, 49, 65*). Further construction of recombinant mutant ScHbIs at Rice University began using the QuickChange site-directed mutagenesis kit in one of three genetic backgrounds: WT, F97Y (high affinity or completely R-state), or K30D (monomeric). Recombinant expression plasmids containing the ScHbI genes were transformed into *E. coli* BL21(DE3) or W3110*lac*Iq L8 cells. All cells were grown in Luria broth (LB; 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl per liter) containing 100 mg/mL

ampicillin. A 20 L Biostat C fermentor was used to express the cells in LB media containing 38 mM δ–ALA. After 4-6 hours of growth at 37 °C, the cells were induced with IPTG, and the temperature was dropped to 28 °C for the remainder of the expression cycle. The fermentor regulated air, pH, and other growth conditions. After an expression period of approximately 15 hours, the cells were harvested by centrifugation and resuspended with lysis buffer (50 mM Tris, pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5 mM dithiothreitol (DTT), 100 mM NaCl, 5 % glycerol, and 1 mM phenylmethanesulfonylfluoride (PMFS)). This modified protocol is based on the method reported by Summerford *et al.* (*46*).

The harvested cell paste was lysed by three freeze thaw cycles, sonication, and passage through a cell breaker two times to disrupt all the cell membranes. $MgCl_2$, DNase I, and RNase A were then added to the liquefied cell paste, and cell debris was separated from the soluble proteins by centrifugation at 30,000 X g (14,000 rpm) for 30 minutes. The supernatant was subjected to two ammonium sulfate precipitations (45% and 95%). The protein in the second precipitate was resuspended and dialyzed overnight against 40 mM 2-(N-cyclohexylamino)ethane sulfonic acid (CHES) buffer, pH 9.0. The isolated protein was then purified by anion exchange chromatography by a fast flow diethylaminoethyl cellulose (DEAE) column at pH 9.0. The purified protein was then concentrated and dialyzed overnight in 40 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), pH 7.0 and passed through a Carboxymethyl (CM) Sepharose size-exclusion column. The purity of the eluted protein by molecular weight was checked by SDS-PAGE, and the spectral characteristics of wt ScHbI are shown in

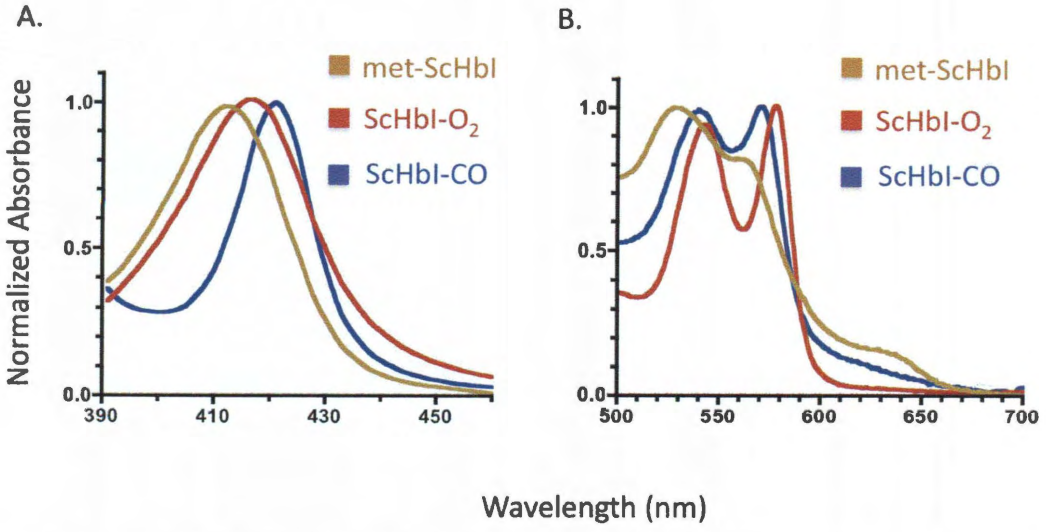


Figure 2.1. Visible spectra for the recombinant *Scapharca inaequivalvis* HbI (ScHbI) expressed and purified at Rice University. *A*, The soret peak for met-ScHbI, ScHbI-O_2, and ScHbI-CO are located at 413 nm, 417, and 421 nm, respectively. *B*, the α and β bands of each protein are labeled accordingly. The absorbance change of each spectrum has been normalized.

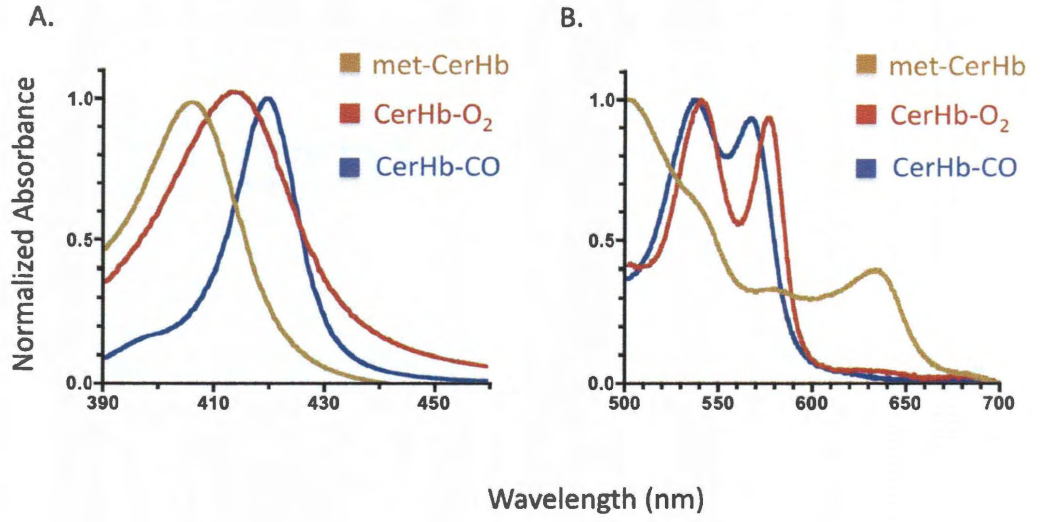


Figure 2.2. Visible spectra for the recombinant *Cerebratulus lacteus* Hb (CerHb) expressed and purified at Rice University. *A*, The soret peak for met-CerHb, CerHb-O_2, and CerHb-CO are located at 406 nm, 413, and 419 nm, respectively. *B*, the α and β bands of each protein are labeled accordingly. The absorbance change of each spectrum has been normalized.

Fig. 2.1. All of the above procedures were conducted using buffers equilibrated with 1 atm CO in order to maintain the stability of the protein. The final purified ScHbI was concentrated to 1 mM, flash frozen, and stored in liquid nitrogen.

Most of the purified recombinant ScHbICO samples were used directly. All experiments were carried out in 0.1 M phosphate buffer, pH 7.0, 1.0 mM EDTA, 20 °C. When necessary, $ScHbIO_2$ protein samples were prepared by equilibration with 1 atm O_2 under an intense light to remove the bound CO. When necessary, the unstable mutants were pre-reduced with dithionite and then quickly passed through a G25 column, and eluted with buffer equilibrated with 1 atm of CO. As noted below, O_2 binding kinetics were often measured in samples containing mixtures of CO and O_2. For those experiments, the starting ScHbI sample was in the reduced CO form, which is extremely stable.

2.3 CerHb sample preparation – Wild type and mutant recombinant CerHb cDNAs were constructed as described previously by the Moen and Dewilde group at the University of Antwerp, using a synthetic gene with codon usage optimized for expression in *E. coli* (*66-68*). Mutants of wt CerHb were then constructed using a QuickChange mutagenesis kit to create a large library of variants. Expression plasmids were transformed into the *E. coli* BL21(DE3)pLysS expression strain. All cells were grown in terrific broth (TB) (12 g of tryptone, 24 g of yeast extract, 1 mL glycerol, 2.3 g KH_2PO_4, and 12.5 g K_2HPO_4 per liter) containing 34 μg/mL chloramphenicol, 100 μg/mL ampicillin, and 1 mM δ-ALA. Four 1L cultures were induced with 0.5 mM IPTG, expressed overnight at 28 °C, and pelleted by centrifugation at 4 °C. Pellets were

resuspended in lysis buffer (0.05 M Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 5 mM DTT) and passed through 3 freeze-thaw cycles (*67*).

Purification began with sonication of the total cell lysate and was followed by two ammonium sulfate precipitations (40% and 90%). The protein was isolated in the last precipitate and was resuspended in and dialyzed against 15 mM Tris-HCl, pH 8.5. The resultant sample was passed through a fast flow DEAE column at low ionic strength and then CerHb was eluted with NaCl. The purified CerHbs were then further purified by 1 or 2 passages through a Sephacryl S200 column (*67*). The purity of the eluted protein was checked by SDS-PAGE, and spectral characterization of wt CerHb is shown in Fig. 2.2. The final purified CerHb was concentrated to 1 mM and stored in liquid nitrogen.

Recombinant CerHb samples were isolated in the reduced oxygenated state and used directly. All experiments were carried out in 0.1 M phosphate buffer, pH 7.0, 1.0 mM EDTA, 20 °C. CerHbCO and CerHbO_2 protein samples were prepared by equilibration with 1 atm CO and 1 atm O_2, respectively. When necessary, the unstable mutants were pre-reduced with dithionite and then quickly passed through a G25 column, and eluted with buffer equilibrated with 1 atm of CO. As noted below, O_2 binding kinetics were often measured in samples containing mixtures of CO and O_2. For those experiments, the starting CerHb sample was in the reduced CO form to prevent autooxidation.

2.4 SwMb sample preparation – The wild type SwMb gene (in pMb413) was originally constructed as described in Springer and Sligar (*69*). Wt SwMb was then altered using a QuickChange site-directed mutagenesis kit to create a number of new single amino acid mutations. Recombinant expression plasmids were transformed into the

E. coli BL21(DE3) expression strain. All cells were grown in LB containing 200 mg/L ampicillin. A 14 L LB culture containing 100-200 μM $FeSO_4$ was expressed overnight at 37 °C with 50 % O_2 tension at pH 7.0. 700 mL of 20 % glucose was added throughout the expression cycle. The cells were pelleted by centrifugation at 4 °C, and the paste was frozen and stored at -20 °C (*69, 70*).

Purification began with lysis of the total cell paste by the addition of lysis buffer (0.5 mM DTT, 1 mM toluene sulfonyl chloride, 40 U/ml DNAse I, 3 U/ml RNAse A, 2 mg/ml lysozyme at pH 6.0), and then incubation overnight at 4 °C. The lysis suspension was pelleted by centrifugation, and subjected to two ammonium sulfate cuts (50% and 90%). The last pellet was resuspended in and dialyzed against high pH buffer (20 mM Tris, 1 mM EDTA, pH 8.0) overnight and passed through an anion exchange column, DEAE 52. The pH was then lowered to 6.0 with 50% acetic acid and the protein was passed over a CM Sepharose size-exclusion column (20 mM NaPi, pH 6.0) and eluted with a pH gradient from 6.0 to 9.5 (*69, 70*). The purity of the eluted protein was checked by SDS-PAGE. Final purified SwMb was concentrated to 1 mM and stored in liquid nitrogen.

Recombinant SwMb samples were isolated in the reduced CO bound state and used directly. All experiments were carried out in 0.1 M phosphate buffer, pH 7.0, 1.0 mM EDTA, 20 °C. SwMbCO and $SwMbO_2$ protein samples were prepared by equilibration with 1 atm CO and 1 atm O_2, respectively. Again, O_2 binding kinetics were often measured in samples containing mixtures of CO and O_2.

2.5 Measurement of overall rates of ligand association and dissociation - CO association time courses were measured after complete laser photolysis of 50 μM CerHbCO, MbCO, and ScHbICO samples containing various free [CO] under pseudo first order conditions. A ~ 500-ns excitation pulse from a Phase-R model 2100B dye laser was used (*71, 72*) and transmittance changes were recorded with a Tektronix TDS220 digitizing oscilloscope. In most cases, the time courses were fitted to a simple exponential expression, and k'_{CO} was calculated as $k_{obs}/[CO]$ or from the slope of plots of k_{obs} versus [CO].

O_2 association and dissociation time courses were determined after complete laser photolysis of 50 μM Hb or Mb samples in mixtures of O_2 and CO using the 500-ns dye laser (*71, 73-75*). In these reactions, various HbCO and MbCO samples are prepared in buffer containing ~ 50 μM CO and 1250 μM O_2, 750 μM CO and 250 μM O_2, 500 μM CO and 500 μM O_2, and 250 μM CO and 750 μM O_2. The laser pulse photolyzes the Fe(II)CO complex, generating deoxyHb or deoxyMb, which can then rapidly react with either ligand. The initial rapid bimolecular rebinding phase is monitored by a large decrease in absorbance of the unliganded Hb or Mb at 436 nm, which is dominated by O_2 rebinding because of its larger association rate constant (Fig. 2.3). The pseudo first order rate for this fast phase, k_{fast}, is given by:

$$k_{fast} = k'_{O2}[O_2] + k_{O2} + k'_{CO}[CO] \qquad \text{(Eq. 2.1)}$$

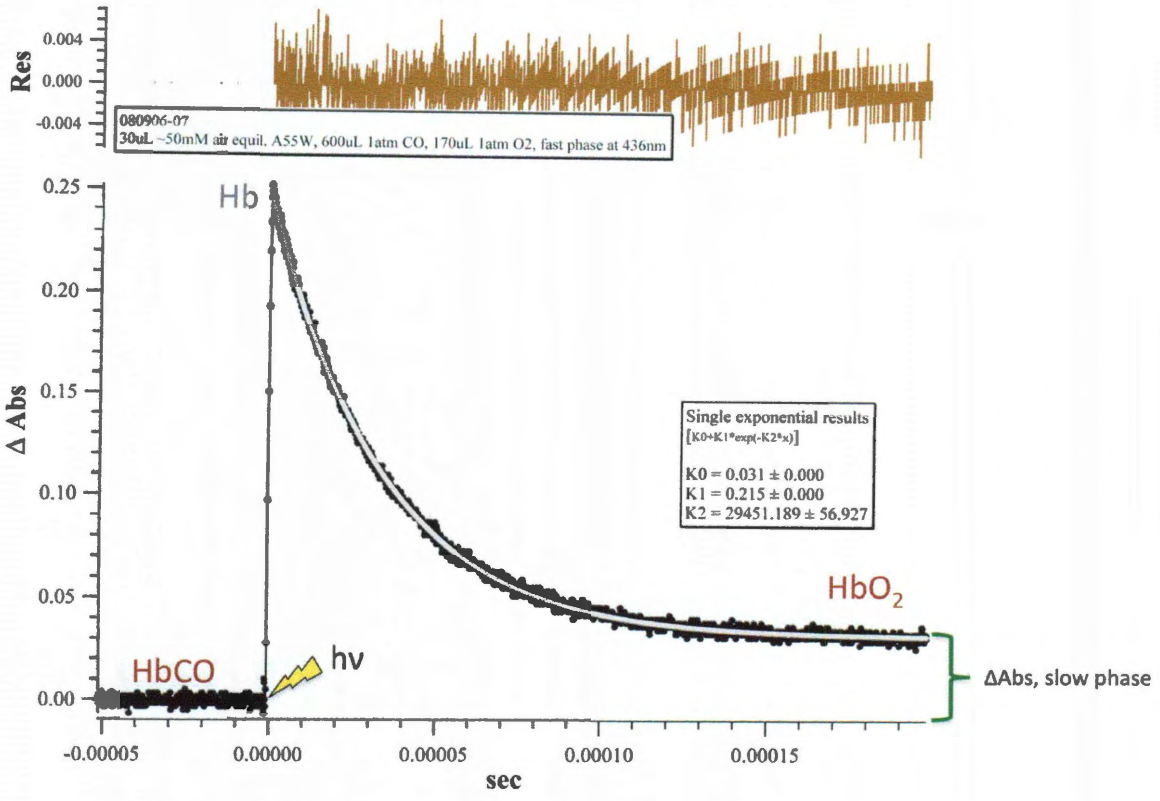


Figure 2.3. Overall ligand binding fast phase time course of CerHb A55W in a 75% CO/ 25% O_2 mixture. The laser excitation pulse photodissociates bound CO causing an almost instantaneous increase in absorbance to form deoxyHb. This rapid increase is followed by a decrease in absorbance at 436 nm, which tracks the fast bimolecular binding of O_2 to the deoxygenated Hb sample.

In this fast phase, CO dissociation can be neglected because the absolute value of k_{CO} is on the order of 0.01 s^{-1}, whereas under the conditions of the laser experiments, k_{fast} is on the order of 10,000 to 300,000 s^{-1}. In most cases, the O_2 dissociation rates, k_{O2}, for the wt and mutant HbO_2 samples of all three proteins (0.2 to ~10,000 s^{-1}) also make only a small contribution to the observed rate coefficients in the fast phase but were considered in the analysis.

The slow phase observed after photolysis of HbCO or MbCO in mixtures of CO and O_2 represents displacement of the transiently bound O_2 by CO, which, although kinetically less reactive, has a higher affinity for the heme Fe atom (*71, 73-75*). The time course for this replacement reaction is best monitored by the large increase in absorbance of HbCO or MbCO at 425 nm (Fig. 2.4). The rate of this slower process, k_{slow}, is given by:

$$k_{slow} = k_{O2}\frac{k'_{CO}[CO]}{k'_{CO}[CO] + k'_{O2}[O_2]} \qquad \text{(Eq. 2.2)}$$

An iterative, non-linear, least squares fitting routine (Solver in Microsoft Excel) was used to optimize the values of k'_{O2}, k_{O2}, and k'_{CO} to give the best fit to the k_{fast} and k_{slow} values at the different $[O_2]/[CO]$ ratios including conditions where $[O_2] = 0$, and where $[CO] \approx 0$ (*71*). In previous work, it has been shown that this method of using photolysis and O_2/CO mixtures gives results identical to those obtained in more conventional rapid mixing and photolysis experiments, and eliminates problems due to autooxidation of unstable mutants and low HbO_2 quantum yields (*73-75*).

In the case of ScHbI samples, partial photolysis time courses were also measured at 436 nm to measure the binding of the second ligand molecule to Hb_2X dimeric

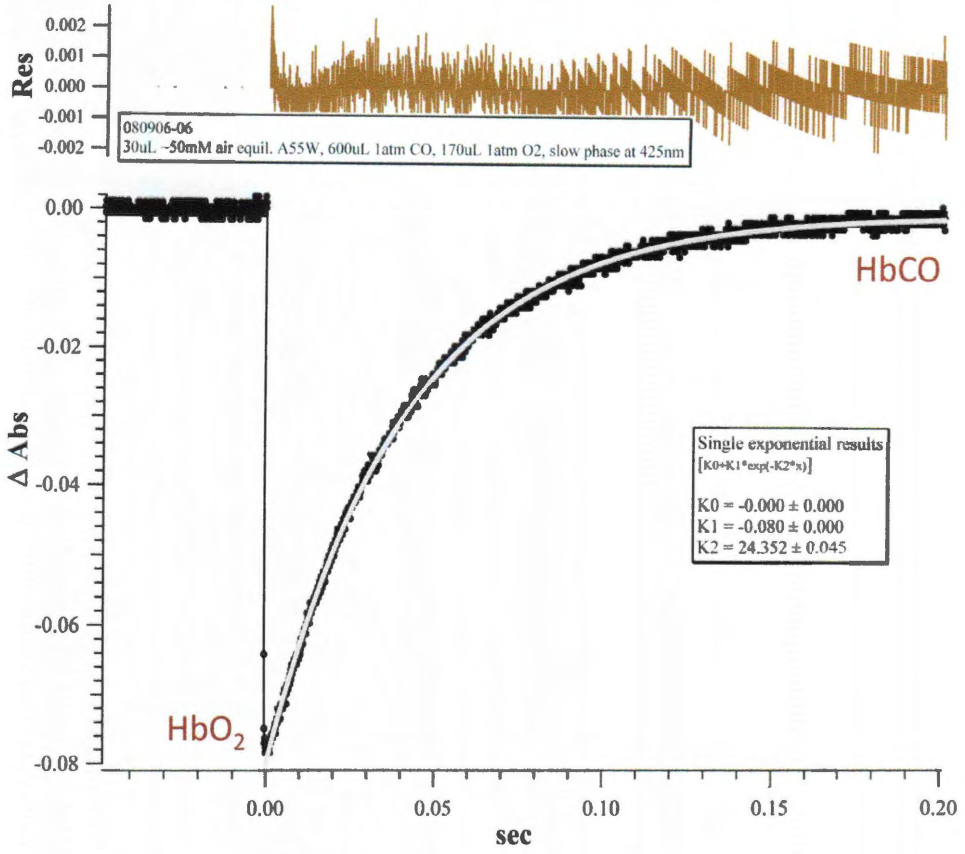


Figure 2.4. Slow O_2 displacement time course of CerHb A55W in a 75% CO/ 25% O_2 mixture. Following photolysis and rapid binding of O_2 to the deoxygenated sample (a small absorbance decrease at 425 nm), transiently bound O_2 is displaced by CO causing a slow absorbance increase at 425 nm.

intermediates. In partial photolysis experiments, the intensity of the exciting light was reduced so that only 10 % of the bound ligands were photolyzed. As a result, the Hb_2X molecules partially switch to the more relaxed, R-state containing one bound ligand, but unfortunately the observed k'_{O2}, k_{O2}, and k'_{CO} rate constants still contained a significant contribution from the T state. Thus in most cases the full and partial photolysis time courses showed similar bimolecular rate constants (see Chapter 3). Again, rate constants for O_2 and CO binding to ScHbI were obtained by fitting the k_{fast} and k_{slow} values to the expressions in Equations 2.1 and 2.2 as was done for CerHb.

The association rate constants for NO binding to deoxy-Hb or deoxy-Mb were measured using a flow-flash multi-mixing apparatus and the 500 ns dye laser. An anaerobic sample of HbCO or MbCO containing little or no free ligand was mixed with 3/4, 1/2, 1/4, and 1/10 atm NO (1500, 1000, 500, and 200 μM, respectively). Immediately after mixing, the HbCO or MbCO complex was photolyzed, and the complete absorbance change for NO binding to the newly generated deoxyHb or deoxyMb was followed at 436 nm. This approach avoids the difficulty of photolyzing the Fe(II)NO complexes, which have an extremely small quantum yield. The observed rate is equal to $k'_{NO}[NO] + k'_{CO}[CO]$. Since both k'_{CO} and [CO] << k'_{NO} and [NO], respectively, this equation reduces to $k_{obs} \approx k'_{NO}[NO]$. The value of k'_{NO} was determined from the slope of k_{obs} *versus* [NO].

CO dissociation time courses were determined by stopped-flow analysis in which the heme bound CO was displaced by a high concentration of the higher affinity NO ligand. In these experiments, 5-10 μM HbCO or MbCO was mixed with 1 atm NO (2000 μM) and the decrease in [HbCO] or [MbCO] was monitored as an absorbance decrease at

420 nm (*72, 76*). Under these conditions, the observed first order rate constant is directly equal to k_{CO} because $k'_{NO}[NO] >> k'_{CO}[CO]$.

2.6 Measurement of rates of geminate recombination- Time courses for internal ligand rebinding (geminate recombination[4]) from within the Hbs and Mbs were measured at 436 nm after excitation with a 9-ns excitation pulse from a Lumonics YAG-laser system. Transmittance changes were recorded on a Tektronix TDS3052 digitizing oscilloscope on nanosecond time scales. The geminate rebinding time courses were fitted to one or two exponential expressions with an offset reflecting the amount of escape to solvent and the fitted exponent equals the rate of geminate recombination, k_{gem} (Fig. 2.5;(*36*)). For a simple exponential internal rebinding process, the fraction of geminate recombination, F_{gem}, was calculated as $F_{gem}= (\Delta A_{gem}/\Delta A_{gem}+Y0)$, where ΔA_{gem} represents the absorbance change associated with internal geminate rebinding (one or two phases). The offset, Y0, is the difference between the absorbance after complete geminate recombination and the absorbance of the original ground state, which is observed either prior to photolysis or at long times after bimolecular rebinding from solvent is complete (*i.e.* fully liganded HbCO or HbO_2) (Fig. 2.5).

2.7 Fourier Transform Infrared (FTIR) spectroscopy – Wild type and mutant HbCO or MbCO samples were exchanged into 100 mM KPi, pH 7.0 buffer and concentrated to ~ 1 mM. 20 μL protein and 200 mM sodium dithionite were added to a CO purged Eppendorf tube, vortexed, and allowed to equilibrate for ≥ 20 minutes. Each 20 μL protein sample was loaded into a BioCell fit with two dry CaF_2 windows. The

[4] Geminate recombination refers to the intramolecular rebinding of a ligand to the heme following photodissociation, but not dissociation from the protein.

spectra were collected using a Nicolet spectrometer with an MCT detector (purged with nitrogen

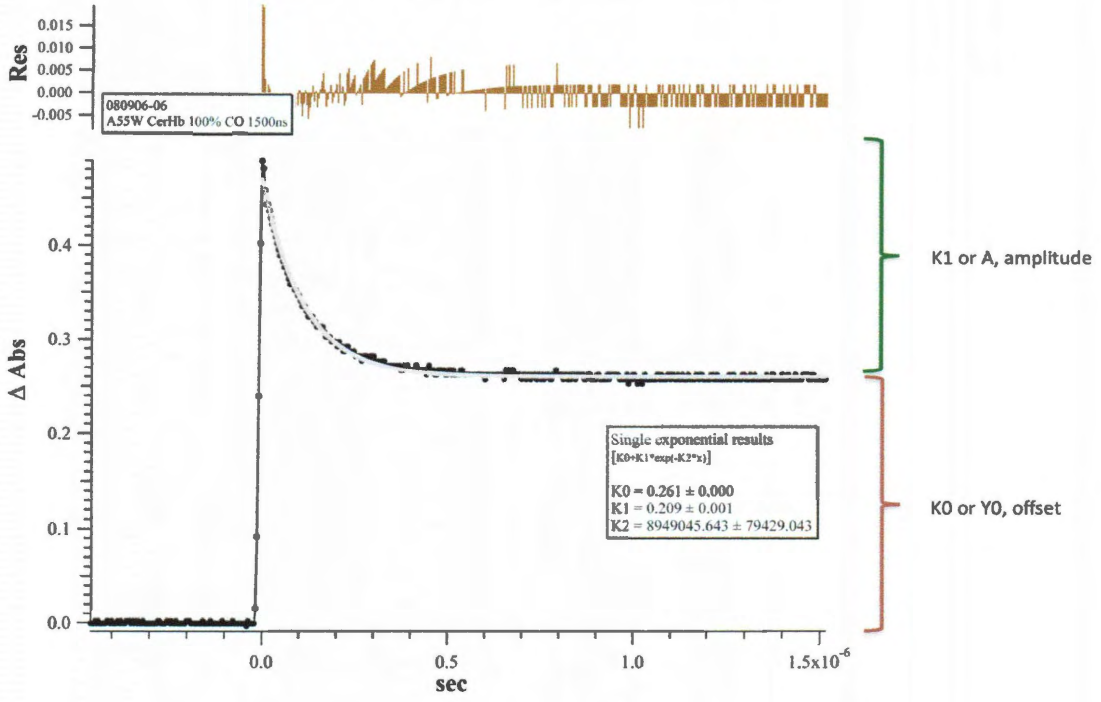


Figure 2.5. Geminate recombination time course of CerHb A55W-CO. Following the 9-ns excitation pulse, a decrease in absorbance is observed at 436 nm. In this example, the amplitude of geminate recombination is 0.209 and the offset is 0.261. Therefore, F_{gem} is 0.44 (F_{gem} = A/ A+Y0). The exponent, K2, is equal to k_{gem}.

≥ 18 hours) and OMNIC software. The protein spectrum was collected between 1800 and 2400 cm^{-1} and subtracted from a background buffer spectrum containing water vapor peaks. Fig. 2.6 displays the spectra generated by the OMNIC software. The process of baseline correction and peak identification was achieved with a program written by George Blouin in Microsoft Excel (*77*). In this program, data between 1900 and 2300 cm^{-1} are used to establish a baseline, and peak integration is achieved to identify major and minor C-O stretching frequencies. Fig. 2.6 provides an example of original corrected spectra generated by this program. Expanded and normalized spectra are given in Fig. 3.3 in Chapter 3. Additional background on FTIR is given in Chapter 4.

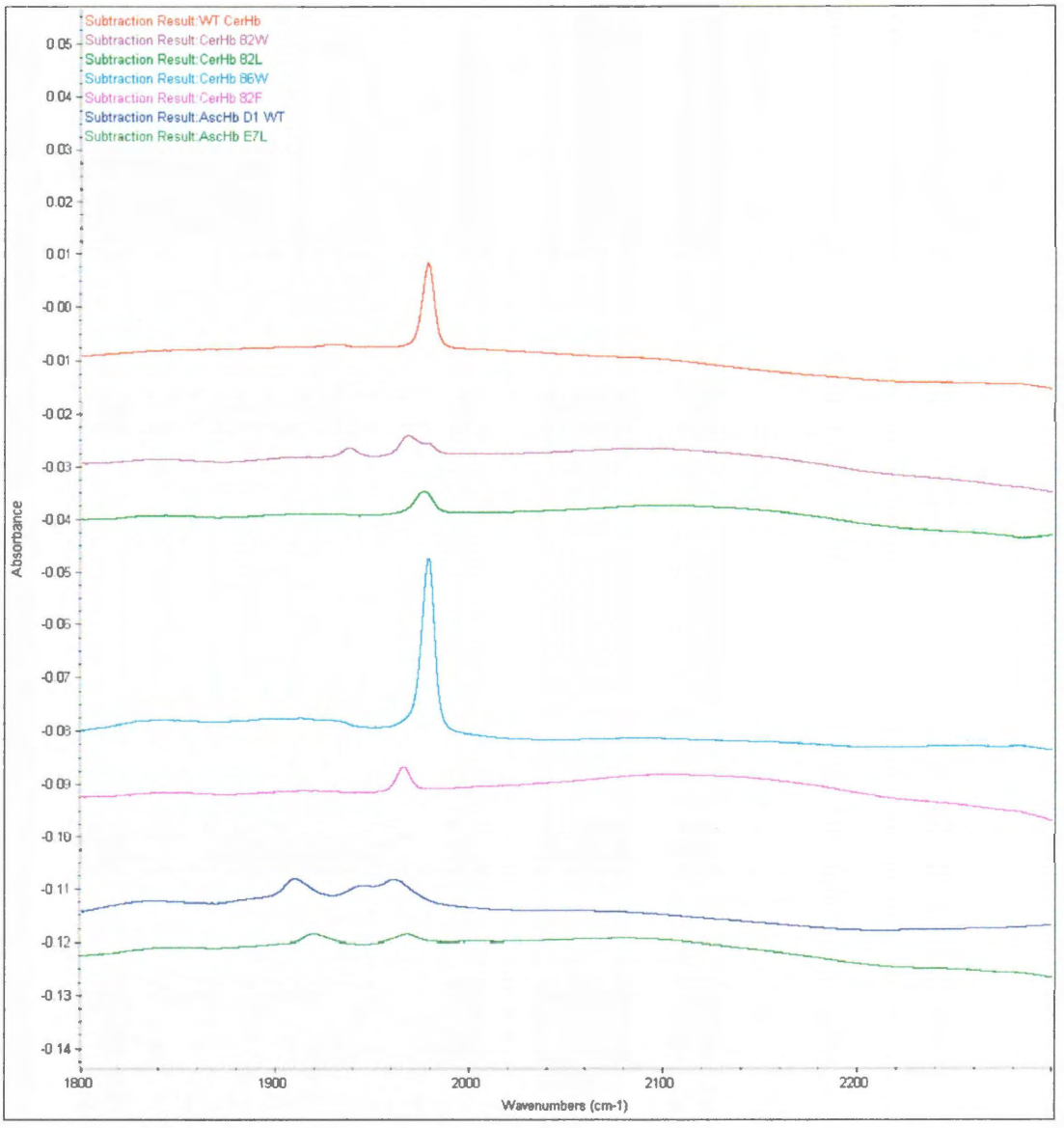


Figure 2.6 FTIR Spectra of heme-bound CO generated with an MCT detector and OMNIC software. The colored lines represent IR absorption spectra (raw data) collected using a Nicolet spectrometer with an MCT detector (purged with nitrogen $\geq$ 18 hours) and OMNIC software. The CO heme-protein spectra are collected between 1800 and 2400 cm^{-1} and subtracted from a background buffer spectrum containing water vapor peaks. Final normalized spectra can be found in Fig. 3.3 in Chapter 3.

Chapter 3: Studies of O_2 and CO Binding to *Scapharca inaequivalvis* Hb (ScHbI): Implications for Quaternary Structure

3.1 Introduction – The overall bimolecular rate at which a small molecule such as CO, O_2 or NO reacts with the Fe of the heme prosthetic group in a globin is governed by the rates at which the ligand enters the protein matrix, moves to the active site, and forms a bond with the heme Fe atom. Thus, it is important to determine the route of ligand migration and the hindrances placed along this pathway. Recently, the ligand diffusion pathways in SwMb and HbA have been examined in detail and, even though changes in quaternary structure do occur upon binding in human hemoglobin, the gated HisE7 pathway first determined for Mb also occurs in both subunits of tetrameric Hb (*36, 73, 78, 79*).

In human HbA, intersubunit contacts are located far away from the solvent exposed edge of the heme group and do not provide any impediment to the outward and upward movement of the distal histidine side chain (Fig. 1.6B). Thus, dissociation of the HbA tetramer into dimers does not affect the accessibility of the E7 channel to solvent in the α and β subunits because the heme propionates and the distal histidine are located far away from either subunit interface. In contrast, many invertebrate hemoglobins, which also cooperatively bind O_2, have a dimeric structure in which the heme propionates are part of the intersubunit contact, and the HisE7 gates appear to be partially blocked by this interface. The best studied example is the dimeric HbI from the blood clam, *Scapharca inaequivalvis* (*44, 45*). In this structure, the intersubunit contacts occur between residues on the E- and F-helices creating an EF:FE dimer interface. This interface appears to have evolved to allow cooperativity in a simple dimer by a direct steric interconnection

between the proximal regions of the heme groups (*51*). The interface between the ScHbI subunits is inverted with respect to the alpha1beta1 interface in HbA, which involves the G and H helices of both subunits. In adition, the EF:EF interface appears to close the E7 gate pathway in which outward movement of the HisE7 side chains would be directed into the dimer interface (Figs. 1.6A, 1.7C, 3.1). As a result, this EF:FE dimer interface could potentially inhibit ligand migration through the HisE7 gates.

Thus, the inverted dimeric hemoglobin I from *Scapharca inaequivalvis* (ScHbI) was chosen as a model for studying a 'closed' E7 gate pathway. However, one key problem is that ligand binding to wt ScHbI is highly cooperative and exists in two conformations, designated R for high O_2 affinity and T for low affinity. This situation complicates kinetic analyses. For example, after complete photolysis, the ScHbI relaxes to the T state in which binding is slower for the first bimolecular association rate constant (i.e. $Hb_2 + X \rightarrow Hb_2X$), whereas CO displacement from the fully oxygenated complex represents ligand dissociation for the fully saturated normally, R state conformation (i.e. $Hb_2X_2 \rightarrow Hb_2X + X$). Even after ~10-15 % partial photolysis to ensure that only Hb_2X molecules are rebinding ligands, the fraction of R state conformers is small. Thus, in our initial screen of ScHbI His69(E7) variants, the association rate constants represent ligand binding to T-like conformations and the dissociation rate constants represent more R-like conformations. In addition, the apparent equilibrium association or affinity constants, defined experimentally as k'_{O2}/k_{O2}, represent hybrid values between the affinities for the R and T states. Thus, it is possible that distal histidine mutations could affect both the quaternary transition and the pathway for ligand binding, resulting in complex changes in the observed rate constants. Despite these complications, the initial screen of the

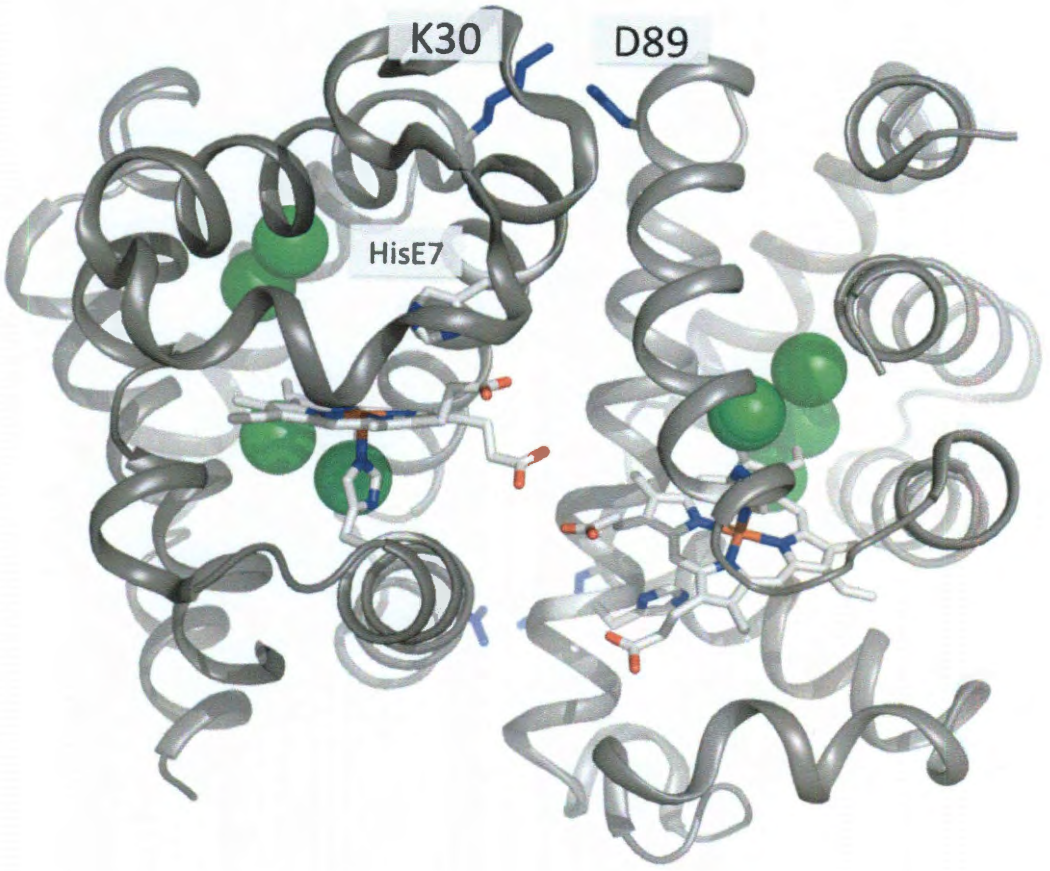


Figure 3.1. Structure of *Scapharca inaequivalvis* hemoglobin indicating the K30:D89 salt bridge and Xe binding sites (PDB ID: 3g46). The K30 and D89 residues (blue sticks) form a key salt bridge in the dimeric ScHbI. The K30D mutation places two negatively charged amino acids, D30 and D89, next to each other, and the charge repulsion causes dissociation into identical monomers. Xe binding sites are designated as green spheres, HisE7 and the heme group are designated as sticks colored by element, and the carbon backbone by gray cartoon. The protein structure was generated in MacPyMol.

His69(E7) mutants still provides key insights into the role of the E7 gate in regulating ligand binding.

However, to avoid the complications due to alternative quaternary conformations, distal histidine mutations in the ScHbI variant with the Phe97(F4)→Tyr mutation were also examined. Normally, PheF4 packs under the heme pocket in T state deoxy ScHbI dimers, and reduces ligand affinity by inhibiting in-plane movement of the heme Fe atom (Figs. 1.7C and 3.2; (*51*)). Insertion of the "extra" hydroxyl group in Tyr97 prevents this packing, and the F4 phenol side chain moves out into the dimer interface, locking the protein in the high affinity, R quaternary state. Thus, when additional mutations are made in the F97Y background, any observed effects are due to changes in the restrictions along the ligand migration pathway in the R state and not changes in the quaternary R to T transition.

Despite the dimer interface, the HisE7 gate appeared to us to be the most plausible ligand migration pathway in ScHbI. However, as in other globins, molecular dynamics (MD) simulations suggest that ligands diffuse away from the heme Fe toward a space defined by residues Ile114, Pro117, Ile118 and Trp135 and then escape from the protein by multiple routes which do not include the HisE7 gate because of the dimer interface (*80*). Recently, Knapp *et al.* (*81*) carried out experiments in the presence of high pressures of Xe gas and showed, by crystallography, that the noble gas occupies some of the cavities suggested by the MD simulations (Fig. 3.1). However, there was little or no effect of high pressures of Xe on ligand binding, suggesting that although internal cavities are able to accommodate Xe, their presence does not affect ligand migration into and out of the active site.

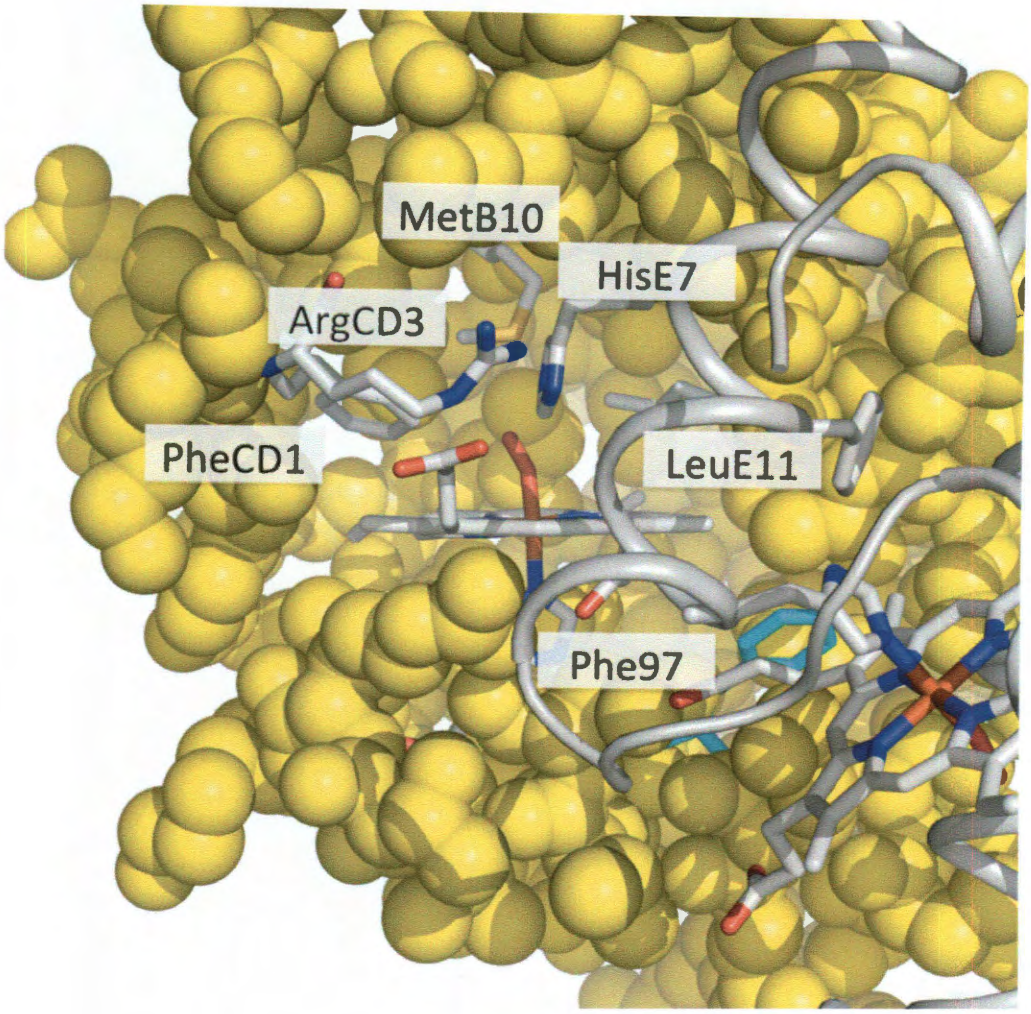


Figure 3.2. Dimer interface structure of *Scapharca inaequivalvis* oxyhemoglobin (PDB ID: 1hbI). The globin backbone of one subunit is represented by yellow spheres and the other subunit by gray cartoon. The key active site residues, MetB10, PheCD1, ArgCD3, HisE7, and LeuE11, are depicted in sticks, and are labeled for the yellow subunit. The Phe97 residue for the yellow subunit is designated by cyan sticks. The perspective of this representation is across the dimer interface where the heme group of the gray subunit is also depicted in sticks. The protein structure was generated in MacPyMol.

To resolve this discrepancy between MD simulations and the Xe-binding experiments and to explore the E7 channel, the distal histidine in ScHbI was mutated systematically to small, intermediate, and large apolar amino acids. These same HisE7 mutations were also constructed in the Phe97(F4)→Tyr genetic background to ensure examination of only the R quaternary structure. Then, as a further control, the Lys30→Asp mutation was also made. This replacement has been shown previously to disrupt the dimer interface causing the formation of monomeric ScHbI (Fig. 3.1;(*49*)). For each series of variants, overall O_2 and CO binding parameters were measured using flash photolysis and rapid-mixing techniques.

3.2 Effects of E7 mutations on FTIR spectra - When HisE7 is replaced by an apolar amino acid, there is a shift in $\nu_{C\text{-}O}$ from 1945 cm^{-1} in wt ScHbI to ~ 1970 cm^{-1} in the Ala, Val, Leu, and PheE7 series (Fig. 3.3, Table 3.1). As discussed further in Section 4.2, there is an inverse relationship between $\nu_{C\text{-}O}$ and the electrostatic field surrounding the bound ligand. Therefore, the shift of ~ 30 cm^{-1} to a higher frequency correlates with a loss of favorable hydrogen bond donation to the bound ligand. When HisE7 is mutated to Gln, a minimal increase in $\nu_{C\text{-}O}$ (~ 5 cm^{-1}) is observed as the amide NH_2 atoms of the Gln side chain can still donate a hydrogen bond to the bound ligand. When HisE7 is mutated to Trp, two $\nu_{C\text{-}O}$ peaks are observed. In this case, the indole side chain appears to have two orientations, a major one with hydrogen bond donation to the bound ligand (low frequency peak) and one with an apolar active site. Two peaks were also reported for TrpE7 SwMbCO (*82*) but only one apolar peak is observed for the TrpE7 mutants of α and β HbA subunits (*73*).

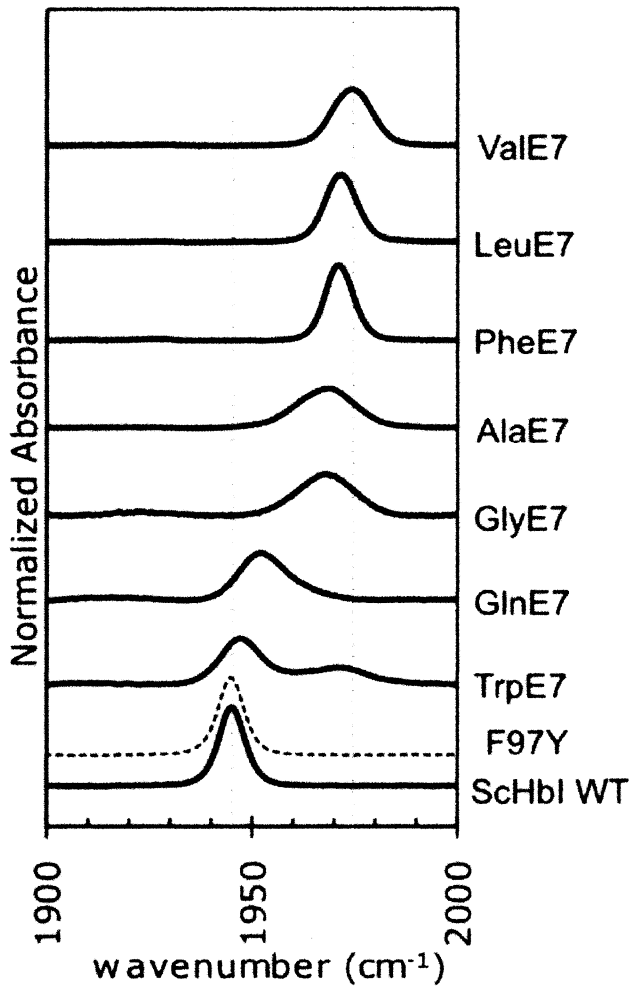


Figure 3.3. FTIR spectra of wt, F97Y, and E7 mutants of ScHbI. These spectra were generated with an MCT detector and OMNIC software followed by baseline correction and peak integration. Wt ScHbI and F97Y ScHbI (dashed) display $\nu_{C\text{-}O}$ peaks at ~1945 cm^{-1}. HisE7 mutations are notated on the right y-axis. As the peak frequency increases from 1945 cm^{-1} to ~1975 cm^{-1}, there is a decrease in hydrogen bond donors surrounding the bound CO (Chapter 4).

Table 3.1. Rate and equilibrium constants for O_2 and CO binding and FTIR stretching frequencies to *A*, position His69(E7) mutants of ScHbI at pH 7.0, 20° C.

WT Background Mutants	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	ν_{C-O} cm^{-1}
Gly69	110	1600	0.017	16	0.060	270	1968
Ala69	63	1900	0.033	13	0.068	190	1969
Val69	73	5300	0.014	27	0.025	1100	1974
Leu69	85	3600	0.024	23	0.017	1350	1972
His69(E7) WT	**14**	**130**	**0.111**	**0.15** **0.24**	**0.005**	**30** **48**	**1945**
Gln69	27	2200	0.012	0.97	0.019	50	1952
Phe69	33	32000	0.001	1.6	0.062	26	1972
Trp69	8	210	0.038	1	0.018	56	1947, 1972

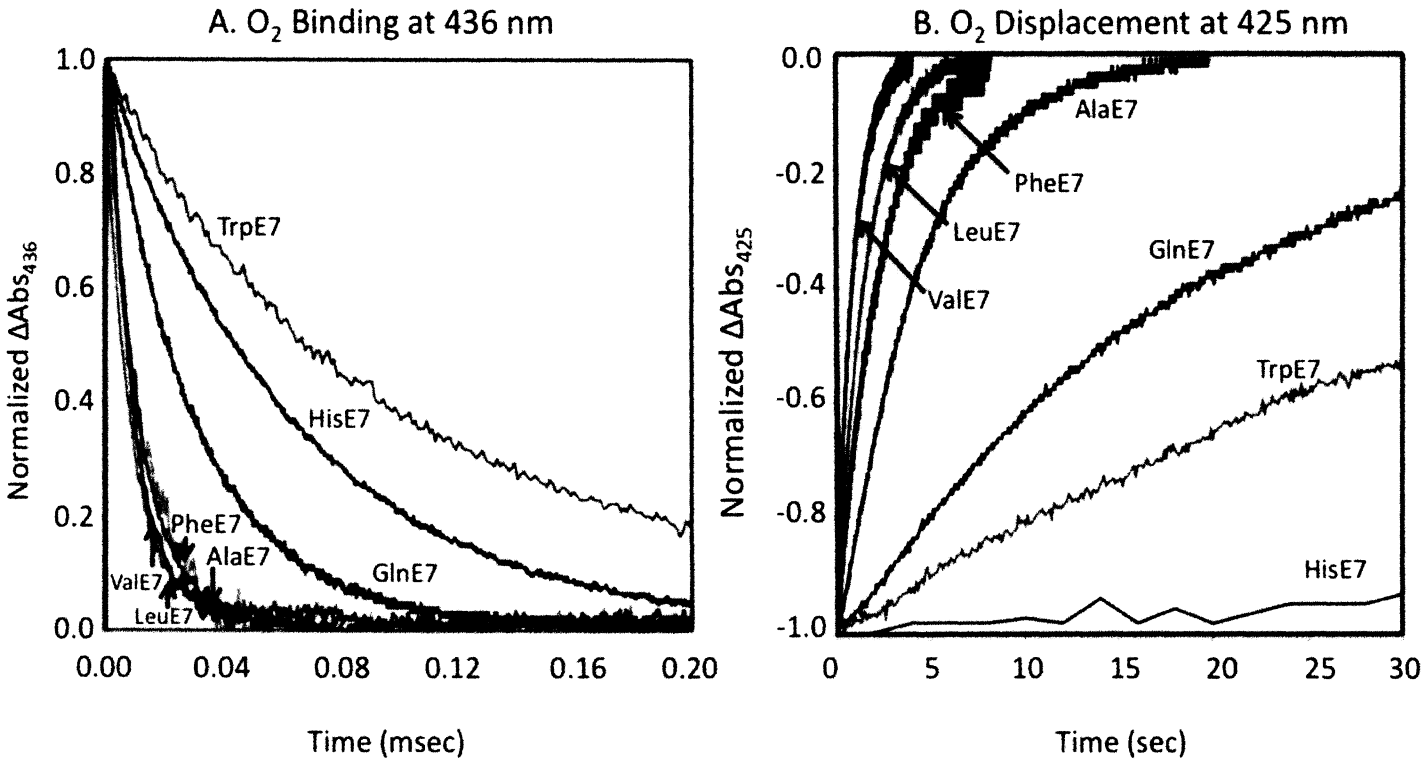


Figure 3.4. Time courses for O_2 binding to wt and HisE7 mutants of ScHbI at pH 7, 20 °C. *A*, bimolecular O_2 rebinding measured at 436 nm after photolysis of 50 μM ScHbI at ~1250 μM O_2 and *B*, Slow O_2 displacement phase measured at 425 nm after photolysis of a mixture of 50 μM ScHbICO, ~500 μM CO and ~500 μM O_2.

A hydrogen bond between the pyrrole NH atoms of the indole ring in TrpE7 and the bound ligand in ScHbICO could account for the $\nu_{C\text{-}O}$ peak at 1947 cm^{-1}, but it is also possible that a water molecule donates a proton when the indole side chain is in a conformation that just blocks but does not enter the E7 channel. The second $\nu_{C\text{-}O}$ peak for TrpE7 at 1972 cm^{-1} suggests that the phenyl portion of the indole side chain is directed toward the bound ligand maintaining an apolar active site (Table 3.1). Such an orientation is observed in the Trp64(E7) SwMbCO crystal structure (PDB ID: 3NML; (*78*)). Unfortunately, a crystal structure of TrpE7 ScHbI was not determined so these interpretations are speculative.

3.3 Effects of E7 mutations on rates of ligand binding to wt ScHbI - When HisE7 is replaced by a small apolar amino acid such as Gly, Ala, Val, or Leu, a 4- 8-fold increase in the bimolecular rate constant for O_2 association occurs (k'_{O2}= 63-110 $\mu M^{-1}s^{-1}$ *versus* k'_{O2} = 14 $\mu M^{-1}s^{-1}$ for wt ScHbI, Table 3.1). Large increases in k'_{O2} are most readily attributed to a decrease in the energy barrier to ligand entry resulting from replacement of the large imidazole gate with smaller amino acids leaving the E7 channel open. In addition, when an apolar amino acid replaces HisE7, distal pocket water is no longer present in the deoxy form and there is no requirement for its displacement prior to ligand binding (K_{H2O}; Fig. 1.4). Both effects will increase the rate constant for bimolecular O_2 binding. The importance of the loss of distal pocket water is illustrated by the larger k'_{O2} for the PheE7 variant, 33 $\mu M^{-1}s^{-1}$ versus 14 $\mu M^{-1}s^{-1}$ for wt ScHbI, where the imidazole side chain in deoxy wt hydrogen bonds to an internal water molecule thus offering more resistance to O_2 binding than the larger benzyl side chain which has no water attached (Table 3.1).

The dependence of the O_2 association rate constant on the size of the E7 amino acid is best examined using apolar mutations in the active site to avoid complications of water binding all together; in this case, any effect is due to steric hindrance alone. The value of k'_{O2} for ScHbI decreases ~ 14-fold from 110 $\mu M^{-1}s^{-1}$ to ~ 8 $\mu M^{-1}s^{-1}$ for the series Gly, Ala, Val, Leu, Phe, and TrpE7 (Table 3.1, Fig. 3.4). The large, but invariant association rate constant for the Ala, Val, Leu series was verified by examining the GlyE7 mutant. Further decreasing the size of the amino acid to Gly resulted in only a 20% increase in k'_{O2}. The same general increase in k'_{O2} with decreasing size at the E7 position is observed for α and β chains of HbA and Mb, supporting the view that the E7 gate pathway occurs in all four proteins (Table 3.2, (*73*)).

Upon replacement of the native HisE7 hydrogen bond donor with an apolar amino acid, the O_2 dissociation rate constant also increases due to loss of stabilization of the bound ligand (i.e. $K_{stabilzation}$ in Fig. 1.4 decreases to 0; Table 3.1). When comparing k_{O2} between the largest and smallest apolar replacements, there is an ~ 20-fold decrease from ~ 4000 s^{-1} for GlyE7 and AlaE7 to 200 s^{-1} in TrpE7 (Table 3.1). However, for this parameter, the dependence on size is complex and probably does not correlate directly with rates of ligand migration out of the protein. For example, replacement of HisE7 with Phe produces a dramatic increase in the O_2 dissociation rate constant to 32,000 s^{-1}, which is 10-fold higher than any of the other apolar mutants. Birukou *et al.* (*73*) observed a similar effect for PheE7 in the α and β subunits of HbA. In all three cases, k_{O2} for the PheE7 mutant is much greater than either AlaE7 or TrpE7 (Tables 3.1 and 3.2). The markedly high dissociation rate constant for PheE7 variants suggest that the benzyl side chain is driven into the distal pocket by a strong hydrophobic effect and sterically hinders

Table 3.2. Rate constants for bimolecular O_2 binding in isolated, monomeric α and β chain mutants at position HisE7 in HbA at pH 7.0, 20° C. Data taken from Birukou *et al.* (*73*).

Subunit	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	Subunit	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}
α Gly(E7)	250	540	0.45	β Gly(E7)	160	680	0.24
α Ala(E7)	160	880	0.18	β Ala(E7)	120	860	0.14
α Leu(E7)	91	450	0.2	β Leu(E7)	110	4600 (40%) 630 (60%)	0.18
α His WT	**40**	**16**	**2.5**	**β His WT**	**52**	**14**	**3.7**
α Phe(E7)	51	9800	0.0052	β Phe(E7)	25	2400	0.01
α Trp(E7)	0.5	450	0.47	β Trp(E7)	1.7	190	0.009

the bound ligand, causing an increase in the rate of thermal bond disruption (k_{-bond} for step 3 in Fig. 1.4).

In contrast to PheE7, the TrpE7 indole ring is almost certainly too big to enter the distal pocket when a ligand is bound and thus probably does not affect the initial ligand dissociation step. However, the TrpE7 indole ring can still occupy a conformation in which it blocks the route of ligand escape following photolysis. This latter interpretation is supported by the decrease in k_{O2} from ~ 4000 s^{-1} in the Gly, Ala, Val, LeuE7 series to 200 s^{-1} in TrpE7 (Table 3.1). This idea is supported by crystal structures of TrpE7 βCO HbA and TrpE7 MbCO, where the indole side chain is found outside of the distal pocket but blocking the E7 channel (*83*).

3.4 Effects of E7 mutations on rates of ligand binding to F97Y ScHbI - Analysis of the kinetic properties of ScHbI variants is complicated by the R to T allosteric transition, which makes it difficult to compare effects of mutations on ligand entry versus changes in quaternary structure. To verify the trends observed for mutants in the wt background, the same HisE7 substitutions were constructed in the R-state (F97Y) genetic background.

As mentioned in Sections 1.5 and 3.1, the larger Tyr97 side chain cannot be accommodated under the heme pocket next to the proximal histidine and is oriented outward toward the dimer interface (Fig. 3.2). As a result, the coordinated HisF8 side chain can adopt a staggered orientation with respect to the pyrrole N atoms of the heme ring, allowing more rapid in-plane movement of the Fe atom during ligand binding (Fig. 1.7D; (*51*)). In globins with an eclipsed orientation of the proximal imidazole ring such as SwMb (*61*), the iron atom is much less reactive because of the steric hindrance to in-

plane movement of the iron atom during ligand binding. Thus, the Fe atom in the F97Y ScHbI mutant is more reactive and its O_2 affinity is 40- 160-fold higher than in wt ScHbI. Since the dimer is locked in the high affinity R state, cooperative ligand binding is eliminated (*51*).

In our control experiments with the single F97Y mutant containing the native HisE7 side chain, two O_2 and CO association and dissociation rates were observed. However, in all cases ligand binding is still very rapid, and the affinities are higher than wt ScHbI. When the distal histidine is mutated within the F97Y genetic background, this kinetic heterogeneity is relieved and simple monophasic time courses for ligand binding and dissociation are observed. For example, upon replacement of His by Val, k'_{O2} increases 2-fold and k_{O2} increases 30-fold, and only single phases are observed (Fig. 3.5, Table 3.3). The key observation among the HisE7 mutations in the F97Y background is that, as in wt ScHbI, small apolar E7 mutations result in significant increases in k'_{O2} from ~ 40 $\mu M^{-1}s^{-1}$ in F97Y to 93 $\mu M^{-1}s^{-1}$ in F97Y/AlaE7 (Table 3.3). When large apolar mutations are constructed as in F97Y/TrpE7, k'_{O2} decreases ~ 6-fold from 93 $\mu M^{-1}s^{-1}$ in F97Y/AlaE7 to 15 $\mu M^{-1}s^{-1}$ in F97Y/TrpE7 (Fig. 3.5). The inverse correlation between the size of the E7 amino acid and the bimolecular rate constant for O_2 still holds, verifying the results for the mutants in the wt background and supporting the E7 channel as the major pathway for ligand entry.

Dramatic increases in k_{O2} also occur when HisE7 is replaced with apolar amino acids, and, again, the largest value of k_{O2} is observed for the HisE7→Phe replacement (3500 s^{-1} for F97Y/PheE7) in comparison to both F97Y (12 s^{-1}) and the mutants containing small apolar amino acids (~ 275 s^{-1}). Again, this marked increase in k_{O2} for the

Table 3.3. Rate and equilibrium constants for O_2 and CO binding to position His69(E7) mutants in the R-state background of ScHbI at pH 7.0, 20° C.

R-State (97Y) Background Mutants	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	ν_{C-O} cm^{-1}
Ala69	93	125	0.74	33	0.004	8,300	
Val69	92	390	0.24	42	0.003	14,000	1974
His69(E7)	**45**	**12**	**3.75**	**2**	**0.005**	**330**	**1945**
97Y	**33**	**2**	**16.5**	**15**	**0.001**	**1,200**	
Gln69	66	63	1.05	20	0.003	6,700	1954
Phe69	56	3500	0.02	11	0.027 0.0001	400 110,000	1970
Trp69	15	57	0.26	2.4	0.007	340	

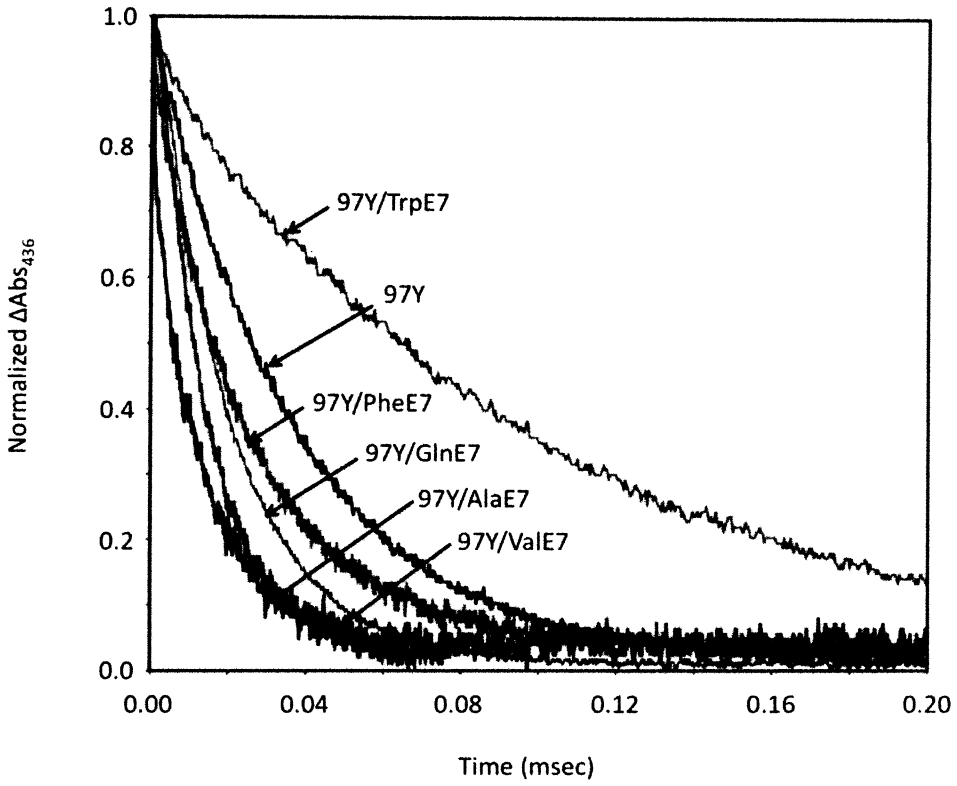


Figure 3.5. Time courses for O_2 binding to 97Y/HisE7 mutants of ScHbI at pH 7.0, 20° C. Bimolecular O_2 rebinding measured at 436 nm after photolysis of 50 μM ScHbICO at ~ 500 μM O_2 and ~ 500 μM CO.

Table 3.4. Rate and equilibrium constants for O_2 and CO binding to position His69(E7) mutants in the monomeric background of ScHbI at pH 7.0, 20° C.

ScHbI	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	ν_{C-O} cm^{-1}
WT	14	130	0.111	0.15 0.24	0.005	30 48	1945
Trp69	8	210	0.038	1.0	0.018	56	
F97Y	45 33	12 2	3.75 16.5	2.0 15	0.005 0.001	330 1,200	1945
F97Y/Trp69	15	57	0.26	2.4	0.007	340	
K30D	43	110	0.41	1.0	0.007	140	1945
K30D/Trp69	13	747	0.02	0.79	0.019	42	

F97Y/PheE7 double mutant is almost certainly due to direct steric hindrance of the bound ligand.

3.5 Kinetics of ligand binding to K30D ScHbI - The results described in the previous sections suggest strongly that the HisE7 residue still functions as a gate for ligand migration in the dimeric ScHbI, despite the presence of the polar interface near the N-terminal region of the E- helix. To test more directly whether or not the dimer interface does retard opening and closing of the E7 gate and ligand migration from the solvent through it, the K30D mutation in ScHbI was constructed which Chiancone *et al.* (*49*) showed causes dissociation into functional monomers. The Lys30(B3) to Asp mutation places a negative charge next to Asp89 at the EF corner of the adjacent subunit in the dimer, which causes complete disruption of the interface (see Fig. 3.1). In these K30D monomers, the HisE7 side chain should be able to rotate more freely into the solvent phase and open the E7 channel (*49*). In addition, this protein should show no cooperative binding, display no apparent kinetic heterogeneity, and be in the high affinity R state.

The single K30D mutation does cause an increase in k'_{O2} from 14 $\mu M^{-1}s^{-1}$ to 44 $\mu M^{-1}s^{-1}$for wt ScHbI (Table 3.4), which could be due to the loss of steric hindrance imposed upon the ligand diffusion pathway by the interface or an increase in Fe reactivity due to the switch to a more high affinity state or both. It appears that the effect of the interface on O_2 binding is negligible when comparing the O_2 binding rates of K30D to F97Y ScHbI. The k'_{O2} values are roughly the same (43 $\mu M^{-1}s^{-1}$ for K30D versus 45 $\mu M^{-1}s^{-1}$ for F97Y ScHbI) for both variants, presumably because they are both in the R state (Table 3.4). In both cases, mutation of HisE7 to Trp causes an identical 3-fold decrease in k'_{O2} to ~ 13 $\mu M^{-1}s^{-1}$ for K30D/TrpE7 and 15 $\mu M^{-1}s^{-1}$ for F97Y/TrpE7 (Table

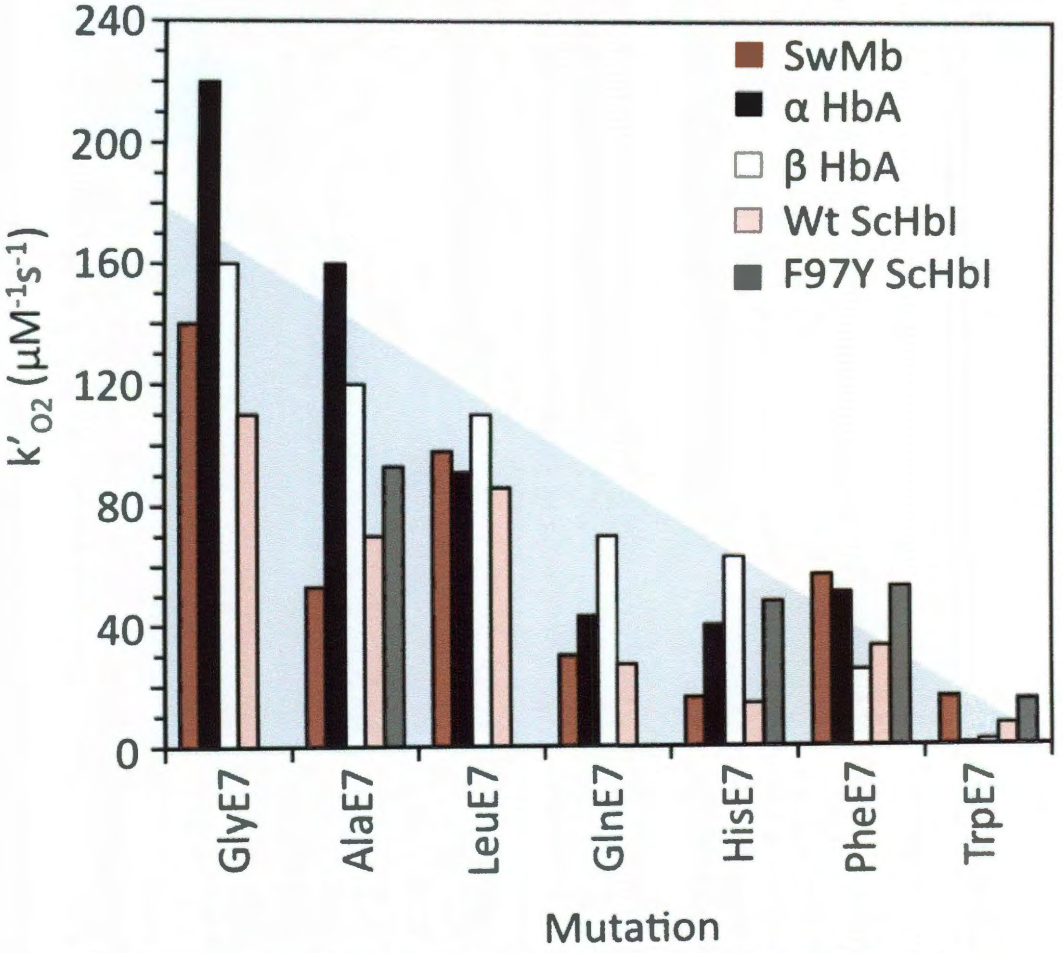


Figure 3.6 Dependence of k'_{O2} on the size of the E7 amino acid side chain for HbA α and β chains, SwMb, and ScHbI E7 mutants in the wt and F97Y genetic backgrounds. There is an inverse correlation between rate of O_2 association and the size of the E7 amino acid side chain for all E7 gate globin models. Previous data for SwMb was taken from (*36*) and all data for HbA was taken from (*73*).

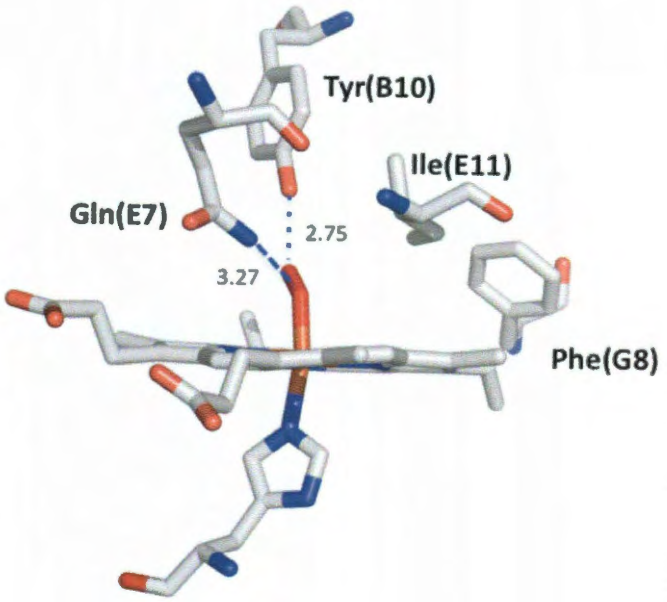
3.4). If the interface had an effect on O_2 binding, one would expect that the O_2 association rates would increase for both K30D and K30D/TrpE7 variants. Thus it seems clear that the dimer interface is not impeding ligand movement through the E7 gate in ScHbI.

3.6 Conclusions - Despite the inverted EF:FE dimer interface in ScHbI, there is still a strong inverse correlation between the size of the amino acid at the E7 position and the bimolecular rate of O_2 association with the protein (Fig. 3.6). Increases in O_2 binding occur upon mutation of the HisE7 to smaller apolar amino acids in both wt and the R state, F97Y genetic background mutations, and a 3-fold decrease in k'_{O2} was observed when HisE7 was replaced with Trp in either wt type dimeric or K30D monomeric ScHbI. As shown in Fig. 3.6, effects of amino acid size at the E7 position in the two dimeric forms of ScHbI are almost identical to those observed previously for the α and β subunits of human HbA tetramers or monomers (*73, 78*) and the SwMb monomer (*36, 79, 84*). Thus, the inverted dimer interface does not appear to change the pathway for ligand entry in ScHbI and probably not in any of the other invertebrate Hbs with this quaternary structure and a distal histidine. To find an alternative pathway for ligand entry, I had to turn to another invertebrate globin with a significantly different active site and globin fold, and the search for that alternative pathway is described in the next two chapters.

Chapter 4: Initial Studies of O_2 and CO Binding to *Cerebratulus lacteus* Hb (CerHb)

4.1 Introduction - As described in Chapters 1 and 3, the distal histidine found in HbA, ScHbI, and SwMb form part of the E7 gate, which can swing open to allow entry of ligands into an interior distal pocket cavity that is circumscribed by the B10, E11 and G8 amino acid side chains. A key question is whether or not alternative pathways occur in other hemoglobins, particularly those without a HisE7 and the classic 3/3 α-helical globin fold. The mini-Hb from *Cerebratulus lacteus* (CerHb) was chosen as a good model system, which has a TyrB10/GlnE7 (YQ) active site motif, an approximate 2/2 α-helical globin fold, and a long apolar tunnel running between the E- and H- helices (*60*). However, I first needed to understand quantitatively how its novel active site affects O_2 and CO binding, regardless of the pathway for entry and exit.

The YQ active site motif is common to both microbial and invertebrate globins (*6, 15*) and characteristically results in an ultra high O_2 affinity. This class of YQ globins is exemplified by *Ascaris suum* hemoglobin, in which TyrB10 and GlnE7 donate two strong hydrogen bonds to bound O_2 resulting in an ultra high affinity (K_{O2} = 575 μM^{-1}) (Fig. 4.1). In contrast, CerHb demonstrates a moderate O_2 affinity (K_{O2} = 1.3 μM^{-1}), which is almost identical to that of SwMb (K_{O2} = 1.1 μM^{-1}) (Fig. 4.2A). As opposed to *Ascaris suum* hemoglobin where the E11 amino acid is an apolar Ile, the E11 position in the distal pocket of CerHb is occupied by a third polar amino acid, Thr. Therefore, rather than two hydrogen bonds between the bound ligand and TyrB10 and GlnE7, additional polar interactions can occur in CerHb between the bound ligand and all three neighboring amino acid side chains (Fig. 4.2B). Surprisingly, the added polarity of the ThrE11 side chain in CerHb reduces O_2 affinity markedly rather than enhancing it, making CerHb an



$K_{O2} = 575\ \mu M^{-1}$
$k'_{O2} = 2.3\ \mu M^{-1}s^{-1}$
$k_{O2} = 0.004\ s^{-1}$

Figure 4.1. Structural representation of *Ascaris* hemoglobin (PDB ID: 1ash) active site. Key amino acids B10, E7, and E11, the heme prosthetic group, and the proximal histidine are depicted in sticks. C atoms are gray, N atoms are blue, and O atoms are red. The two strong hydrogen bonds are depicted by blue dashed lines with the approximate distances (in Å) labeled. The figure was generated in MacPyMol.

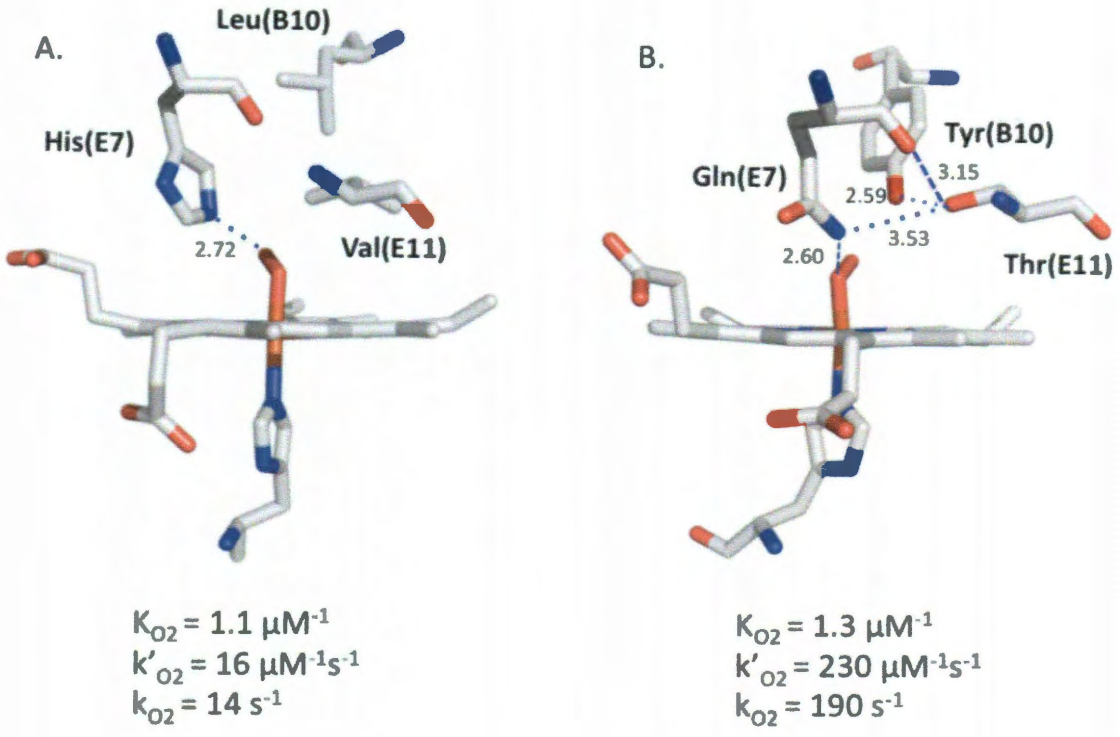


Figure 4.2. Structural representation of *A*, SwMb (PDB ID: 2mgm) and *B*, CerHb (PDB ID: 1kr7) active sites. Key amino acids B10, E7, and E11, the heme prosthetic group, and the proximal histidine are depicted in sticks. C atoms are gray, N atoms are blue, and O atoms are red. In each panel, the relevant hydrogen bonds are depicted by blue dashed lines with the approximate distances (in Å) labeled. The figure was generated in MacPyMol.

unusual YQ Hb with a moderate O_2 affinity. A structural explanation of this moderate affinity was obtained by a combination of mutagenic, crystallographic, and rapid kinetic measurements (See Section 4.6; (*42, 60, 62*)).

The effects of polarity in the active site in CerHb are unique among YQ motif containing globins. (1) The favorable electrostatic interactions between ThrE11 and TyrB10 and GlnE7 in CerHb led to an unfavorable interaction with bound O_2 and a marked decrease in affinity. (2) The amide side chain of GlnE7 in CerHb is "tied" down by 3- 4 hydrogen bonds to amino acids surrounding the heme-binding site, perhaps inhibiting entry and exit through the E7 gate.

4.2 Correlations between $\nu_{C\text{-}O}$, electrostatic field, k_{CO} and k_{O2} in SwMb - The peak frequency for C-O bond stretches, $\nu_{C\text{-}O}$, is directly related to bond order. A higher bond order results in a higher stretching frequency peak ($\nu_{CO} \approx 2150$ cm^{-1} for a triple bond) and a lower bond order results in a smaller frequency peak ($\nu_{CO} \approx 1650$ cm^{-1} for a double bond). The same principles apply when the CO molecule is bound to the heme Fe atom. When CO is bound to Hb or Mb, it exhibits a bond order of 2.5-2.8 which leads to strong absorption of IR light in a region distinct from the rest of the protein matrix and water, ~1900 to 2000 cm^{-1} (*85*). The precise bond order is dependent upon the surrounding electrostatic field, which is determined by the amino acid side chains and their orientation in the distal portion of the heme pocket. Thus, residues B10, E7, and E11 can affect bond order that, in turn, determines the peak stretching frequencies in the IR absorption spectrum. As a result, FTIR measurements of iron carbonyl complexes have been used as an empirical predictor of the electrostatic potential field surrounding bound ligands in Mbs, Hbs, and other heme-containing proteins (*86-88*).

In the case of Mb, a hydrogen bond exists between the bound CO and the amide Nε-H atoms of HisE7; the C-O bond order is ~ 2.6; and the peak $\nu_{C\text{-}O}$ is located at 1945 cm^{-1} (Fig. 4.3B). Increasing the positive field adjacent to the bound ligand causes the C-O bond order to decrease and the Fe-C bond order to increase due to stabilization of the more carbonyl like resonance structure (*i.e.* Fe=C=O) with more non-bonded electron density on the O atom and greater back bonding by the Fe atom (Fig. 4.3A). Removal of positive field or hydrogen bond donation to create an apolar active site increases the C-O bond order, and addition of negative field from Lewis bases (electron donors, normally non-bonded electrons on O or N atoms) cause even further increases due to stabilization of the $Fe^{\delta-}$-C≡$O^{\delta+}$ resonance form (Fig. 4.3C; (*85*)). In the latter cases, $\nu_{C\text{-}O}$ increases and $\nu_{Fe\text{-}C}$ decreases. The net result is an inverse relationship exists between the electrostatic field near the bound ligand and the peak C-O stretching frequency, ν_{CO} (*33, 85*).

Even though the electrostatic environment surrounding bound CO does affect the C-O and Fe-C bond orders, the magnitudes of these effects are small in terms of changes in the CO dissociation rate and association equilibrium constants because the Fe-CO complex is apolar with little or no net displacement of charge from the Fe atom. In contrast, the Fe-O_2 complex is highly polarizable and resembles a ferric-superoxide complex, i.e. Fe(III)-$O_2^{\bullet-}$, with a partial charge of -0.8 on the O atoms (*85*). As a result, hydrogen bond donation or an adjacent partial positive charge from a Lewis acid can stabilize bound O_2, decreasing the rate of dissociation several orders of magnitude (*85, 89*). Thus, a strong linear correlation is expected between the C-O stretching frequency of bound CO, which measures the electrostatic field, and the logarithm of the O_2

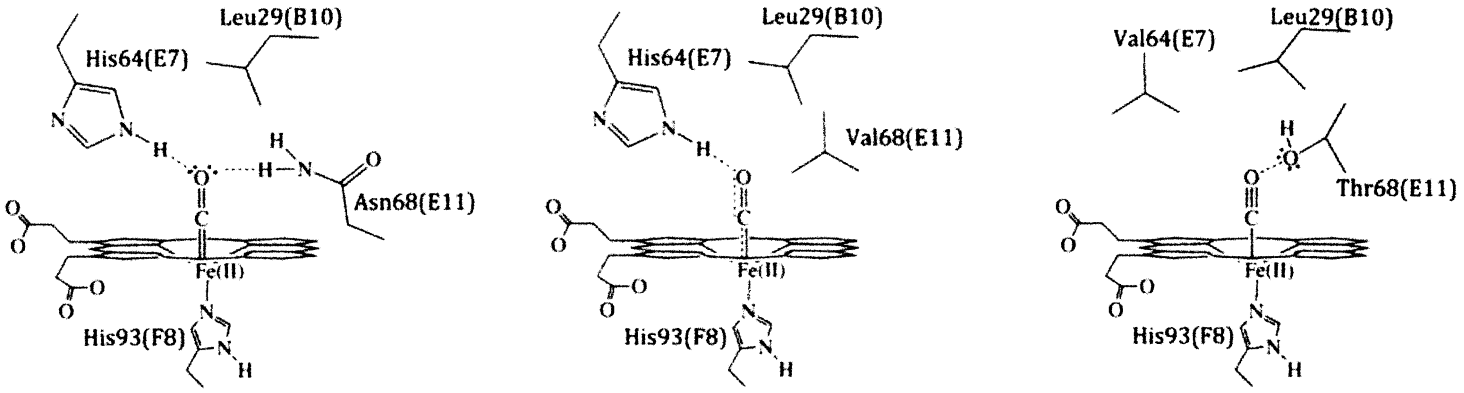


Figure 4.3. Proposed models for electrostatic interactions in the active site of SwMb. *A*, Val68(E11)Asn demonstrates a positive electrostatic field with two hydrogen bond donors surrounding the bound CO. The Fe=C=O resonance form dominates and a low ν_{C-O} of 1916 cm^{-1} is observed. *B*, Wt SwMb demonstrates a more neutral electrostatic field with one hydrogen bond donor surrounding the bound CO. As a result, a resonance form where the C-O bond order is ~2.6 dominates and a ν_{C-O} of 1945 cm^{-1} is observed. *C*, His64(E7)Val/Val68(E11)Thr SwMb demonstrates a negative electrostatic field where both hydrogen bond donors have been removed. The non-bonded electrons of Thr68(E11) are directed towards the bound CO. The Fe-$\bar{C}\equiv O^{+}$ resonance form dominates and a higher ν_{C-O} of 1984 cm^{-1} is observed (*62*). Structures generated with ChemDraw. Figure taken from (*85*).

dissociation rate constant from the same globin.

As shown in Fig. 4.4[5], this correlation has been established unambiguously using O_2 dissociation rate constants and C-O stretching frequencies from a large library of SwMb variants (*33, 82*). This correlation is demonstrated most dramatically by comparing wt SwMb (1945 cm^{-1}) to both the single mutant Val68(E11)Asn and the double mutant His64(E7)Val/Val68(E11)Thr (Table 4.1A, Fig. 4.3). As described by Phillips (*85*), mutation of ValE11→ Asn in the single mutant provides additional hydrogen bonding to the bound ligand resulting in a more positive electrostatic field surrounding the bound ligand (Fig. 4.3A). On the other hand, mutation of HisE7→Val and ValE11→Thr results in the non-bonded electrons of the ThrE11 hydroxyl O atom being directed towards the bound CO atom yielding a more negative electrostatic field (Fig. 4.3C). In the case of the single ValE11Asn replacement, the neutral Fe=C=O resonance form is more favored by the "extra" hydrogen bond and a lower $\nu_{C\text{-}O}$ of 1916 cm^{-1} is exhibited in comparison to 1945 cm^{-1} in wt SwMb. In the case of the double mutant, the $Fe\text{-}^{-}C{\equiv}O^{+}$ resonance form is favored and a higher $\nu_{C\text{-}O}$ of 1984 cm^{-1} is exhibited ((*85*), Fig. 4.3, Table 4.1A).

4.3 Construction of SwMb vs CerHb single and multiple mutants - In order to gain a better understanding of the electrostatic environment generated by the B10, E7, and E11 side chains that surround bound ligands in both proteins, SwMb mutants that mimic the distal pocket of CerHb and CerHb mutants that mimic SwMb were constructed and analyzed. The single SwMb mutants HisE7Gln and ValE11Thr, the double

[5] Data from α and β subunits of HbA (73) and F97Y mutants of ScHbI (Chapter 3) have also been included in this data set to demonstrate the more general correlation between $\nu_{C\text{-}O}$ and $\log(k_{O2})$ among 3/3 α-helical globins with a distal histidine.

Table 4.1. CO stretching frequency, O_2 dissociation rate, and O_2 affinity of *A*, key SwMb mutants and wt *Ascaris suum* Hb, and *B*, key CerHb mutants. Data taken from (*36, 61, 85, 89-91*).

A. SwMb Mutants	$\nu_{C\text{-}O}$ (cm^{-1})	k_{O2} (s^{-1})	*log* k_{O2}	K_{O2} (μM^{-1})
WT	1945	15	1.2	1.1
Single				
Leu29(B10) Phe29	1932	1.4	0.15	15
His64(E7)				
Leu64	1965	4100	3.6	0.02
Gln64	1945	130	2.1	0.01
Val68(E11)				
Asn68	1916	0.54	-0.3	3.5
Thr68	1958	39	1.6	0.08
Multiple				
Leu29(B10)Tyr/ His64(E7)Gln	1934	1.6	0.2	1.8
His64(E7)Val/ Val68(E11)Thr	1984	4000	3.6	0.025
Leu29(B10)Tyr/ His64(E7)Gln/ Val68(E11)Thr	1938, 1953	7.4	0.9	0.06
Ascaris suum Hb D1	1912	0.004	-2.4	370

Table 4.1 (cont'd). CO stretching frequency, O_2 dissociation rate, and O_2 affinity of *A*, key SwMb mutants and wt *Ascaris suum* Hb, and *B*, key CerHb mutants.

B. CerHb Mutants	$\nu_{C\text{-}O}$ (cm^{-1})	k_{O2} (s^{-1})	*log* k_{O2}	K_{O2} (μM^{-1})
WT	1979	190±25	2.3	1.3
Single				
Val7(B6)				
Ala7	1981	270	2.4	0.93
Trp7	1974	14	1.1	2.4
Phe10(B9)				
Ala10	1971	40	1.6	1.8
Tyr11(B10)				
Ala11	1976	400	2.6	0.30
Phe11	1969	460	2.7	0.30
Gln44(E7)				
Leu44	1993	96	2.0	1.9
Phe44	1987	170	2.2	0.12
His44	1962	81	1.9	1.0
Lys47(E10)				
Trp47	1983	180	2.3	1.0
Thr48(E11)				
Ala48	1934	0.26	-0.59	160
Val48	1940	0.025	-1.6	170
Ala82(G8)				
Leu82	1980	56	1.7	0.88
Phe82	1965	18	1.3	13
Trp82	1960	11	1.0	5.7
Leu86(G12)				
Trp86	1980	60	1.8	1.6
Multiple				
Y11F/Q44L[a]	1972	490	2.7	0.60
Y11F/Q44H	1941	16	1.2	0.88
Y11F/T48V	1951	62	1.8	3.7
Q44L/T48V	1940	4.2	0.62	13
Y11F/Q44L/T48V	1968	270	2.4	1.6
Y11F/Q44H/T48V	1935	2.9	0.46	10

[a] This group of mutants follows the conventional single letter mutation nomenclature where Y11F means mutation of Tyr to Phe at the 11th amino acid from the N-terminus.

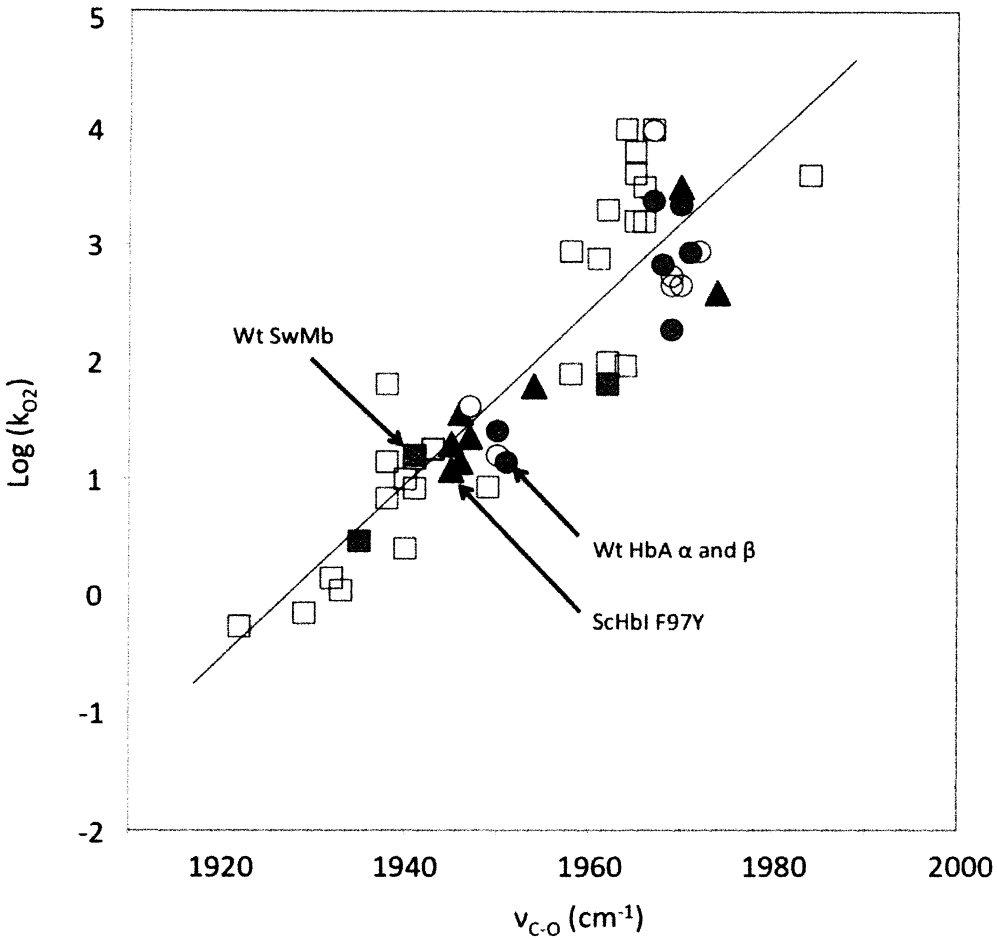


Figure 4.4. Correlations between the C-O stretching frequency (ν_{C-O}) and O_2 dissociation rate constants (k_{O2}) in SwMb, HbA, and ScHbI. A direct relationship exits between ν_{C-O} and log(k_{O2}) for SwMb(*92, 93*), HbA (*73*) and ScHbI F97Y mutants (Chapter 3). SwMb single mutants are represented by open squares and CerHb mimics of SwMb are represented by filled squares. HbA α subunit mutants are represented by open circles, and HbA β subunit mutants are represented by filled circles. ScHbI F97Y mutants are represented by filled triangles. Data for wt ScHbI were not included because the k_{O2} values represent a mixture of R and T values and do not correspond directly to a single quaternary state. The straight line fit has an R-square value = 0.78756 .

mutant LeuB10Tyr/HisE7Gln, and the triple mutant LeuB10Tyr/HisE7Gln/ValE11Thr had been constructed previously, and I fully characterized the FTIR spectra of all four variants (Table 4.1). The CerHb single mutants, GlnE7His and ThrE11Val had also been constructed previously. To complete the direct comparison between CerHb and SwMb, the multiple CerHb mutants GlnE7His/ThrE11Val and TyrB10Phe/GlnE7His/ThrE11Val were generated.

4.4 Measurement of O_2 binding for active site mutants in SwMb - The effects on O_2 and CO binding in some of the SwMb mutants that mimic the distal pocket of CerHb have been reported previously (*53, 61, 62, 85, 90*). Mutation of the HisE7 to Gln in SwMb causes an ~ 10-fold increase in k_{O2} and a 6-fold decrease in K_{O2} (Table 4.1) due to weakening the hydrogen bond between the E7 side chain and bound O_2. Relative to the single GlnE7 SwMb mutant, the double SwMb mutant LeuB10Tyr/HisE7Gln restores k_{O2} and K_{O2} to wt SwMb values due to an additional hydrogen bond between the TyrB10 hydroxyl group and bound O_2. As part of her Master's Thesis, Angela Hale created the triple SwMb mutant, LeuB10Tyr/HisE7Gln/ValE11Thr, which mimics the CerHb distal pocket. She found that the addition of ThrE11 yields an increase in k_{O2} and a decrease in K_{O2} in relation to the double TyrB10/GlnE7 SwMb mutant. Thus, as in wt CerHb, the ThrE11 hydroxyl atom in the SwMb triple mutant pulls the proton on the TyrB10 hydroxyl away from bound O_2 destabilizing it, verifying our previous interpretations ((*62, 90*), Table 4.1).

4.5 Measurement of O_2 binding and affinity for active site mutants in CerHb - Pesce *et al.* (*62*) constructed the GlnE7→ His CerHb mutant and reported decreases in both k'_{O2} and k_{O2} and a small increase in K_{O2} in comparison to wt CerHb, the latter being

Table 4.2. O_2 association and dissociation rate constants, O_2 affinity, and CO dissociation rate constants of *A*, key SwMb mutants and *B*, key CerHb mutants. Data taken from [i] (*85*), [ii] (*70*), [iii] (*61*), [iv] (*90*), [v] (*62*), and [vi] (*53*).

A. SwMb Mutant	k'_{O2} ($\mu M^{-1}s^{-1}$)	k_{O2} (s^{-1})	K_{O2} (μM^{-1})	k_{CO} (s^{-1})
WT	17	15	1.1	0.019
His(E7)Gln[i]	23	130	0.18	0.012
Leu(B10)Phe[ii]	21	1.4	15	0.006
Val(E11)Thr[i]	2.8	39	0.072	0.079
Leu(B10)Tyr/ His(E7)Gln[iii]	2.8	1.6	1.8	0.019
Leu(B10)Tyr/His(E7)Gln/ Val(E11)Thr[v]	0.45	7.4	0.061	0.021

B. CerHb Mutant	k'_{O2} ($\mu M^{-1}s^{-1}$)	k_{O2} (s^{-1})	K_{O2} (μM^{-1})	k_{CO} (s^{-1})
WT	230	190	1.2	0.054
Gln(E7)His	130[v]	65[v]	1.9[v]	0.011[vi]
Tyr(B10)Phe	140	460	0.30	0.0081
Thr(E11)Val[v]	30	0.18	170	0.0070
Gln(E7)His/ Thr(E11)Val	120	11	11	0.010
Tyr(B10)Phe/Gln(E7)His/ Thr(E11)Val	29	2.9	10	0.036, 0.0050

due to a slightly stronger hydrogen bond between the HisE7 side chain and bound O_2 (Table 4.2) . To expand the comparison between the SwMb and CerHb active sites, the double CerHb mutant GlnE7His/ThrE11Val was constructed. By replacing the ThrE11 with Val, a strong hydrogen bond can now form between TyrB10 and bound O_2 as observed for the single ValE11 mutant, but the effect in the presence of HisE7 is much smaller, suggesting competition between HisE7 and TyrB10 for favorable interactions with the bound ligand. A similar competition occurs in the single TyrB10 SwMb mutant where HisE7 is present naturally (*61*). Replacing TyrB10 with Phe in HisE7/ValE11 CerHb leads to smaller rates of O_2 binding and release with little change in affinity (Table 4.2). The O_2 affinity and rate constants of this PheB10/HisE7/ValE11 CerHb triple mutant are roughly equal to those of the PheB10 single SwMb mutant, demonstrating that site directed mutagenesis in CerHb can produce a mutant that mimics the corresponding variant in myoglobin.

4.6 Electrostatic stabilization in CerHb - In the case of wt CerHb, the non-bonded electrons of the TyrB10 hydroxyl O atom point toward the bound CO stabilizing the $Fe^{-}C{\equiv}O^{+}$ resonance form (Fig. 4.5A), an effect similar to that observed for the double HisE7Val/ValE11Thr mutation in SwMb (Fig. 4.3C). As a result, there is a negative electrostatic field surrounding the bound ligand; a high $\nu_{C\text{-}O}$ peak occurs at 1979 cm^{-1}, and rapid rates of O_2 dissociation are observed for wt CerHb (Table 4.1B). By mutating the CerHb polar active site residue ThrE11 to a Val, the TyrB10 phenol H atom rotates toward the bound O_2, stabilizing the Fe=C=O resonance structure, an effect similar to that observed in wt and Val(E11)Asn SwMb mutant (Fig. 4.3B). As a result, the CO stretching frequency decreases dramatically from 1979 cm^{-1} in wt CerHb to 1940 cm^{-1} in

the ValE11 CerHb mutant and a dramatic 1000-fold decrease in k_{O2} is observed (Tables 4.1 and 4.2). As a result, there is a 140-fold increase in O_2 affinity (K_{O2} increases from 1.2 μM^{-1} in wt CerHb to 170 μM^{-1} in ValE11 CerHb (*62*)). Thus, this isosteric replacement of Thr by Val markedly alters the hydrogen bond network between ThrE11, TyrB10, and GlnE7, and allows strong hydrogen bonds to form between the bound O_2 atoms and the side chains of TyrB10 and GlnE7. These strong hydrogen bonds lead to the high K_{O2}, which is characteristic of other invertebrate hemoglobins with the YQ active site motif (Fig. 4.5B, Table 4.1B, (*6, 15*), (*62*).

When all three active site residues in CerHb are mutated to apolar amino acids in the Tyr11B10Phe/Gln44E7Leu/Thr48E11Val CerHb mutant, the electrostatic field becomes zero and a decrease in C-O stretching frequency is observed, from 1979 cm^{-1} in wt CerHb with a negative field from TyrB10 to 1968 cm^{-1} in the completely apolar CerHb triple mutant (Table 4.1C). The neutral electrostatic field in the Tyr11B10Phe/Gln44E7Leu/Thr48E11Val CerHb mutant resembles that in the HisE7Leu SwMb mutant, which shows a similar ν_{C-O} peak at 1967 cm^{-1} (Table 4.1A). When the B10 and E11 active site residues are apolar, but the E7 residue is mutated to His as in the Tyr11B10Phe/Gln44E7His/Thr48E11Val CerHb mutant, the positive increase in electrostatic field causes a decrease in ν_{CO} to 1935 cm^{-1}, a stretching frequency which is nearly equal to that observed in the PheB10 SwMb mutant (1932 cm^{-1}) (Table 4.1C). This ~ 35 cm^{-1} decrease in ν_{C-O} for the triple Tyr11B10Phe/Gln44E7His/Thr48E11Val CerHb mutant in comparison to wt CerHb is accompanied by a marked 65-fold decrease in k_{O2} (Table 4.1 and Fig. 4.6). Fig. 4.6 demonstrates that correlations between ν_{C-O} and k_{O2} also occur for the library of CerHb mutants and related globins, and further verifies using

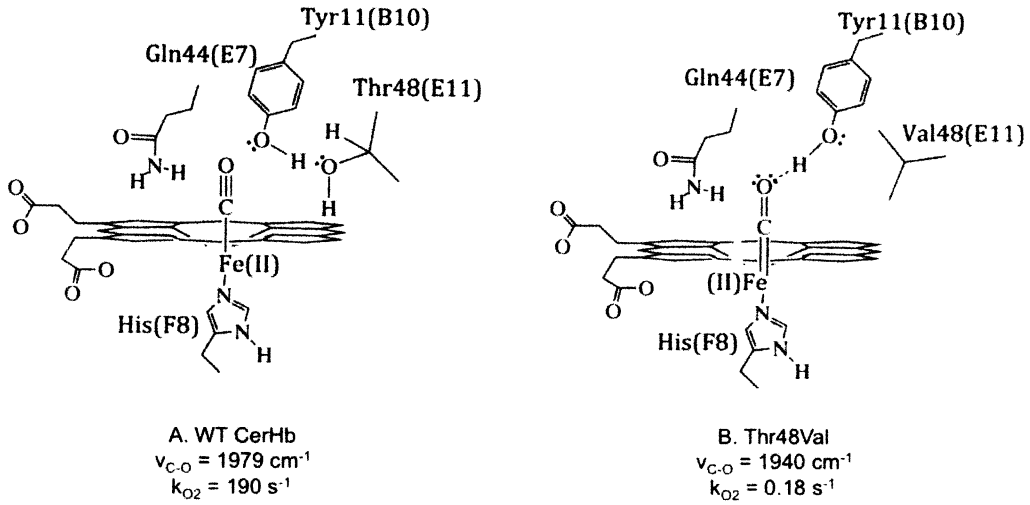


Figure 4.5. Proposed models for electrostatic interactions in the active site of CerHb. *A*, Wt CerHb demonstrates a negative electrostatic field surrounding the bound CO as a result of the TyrB10 non-bonded electrons directed towards the bound CO. The Fe-$^{-}$C≡O^{+} resonance form dominates and a high ν_{C-O} of 1979 cm^{-1} is observed. *B*, Thr48(E11)Val CerHb demonstrates a positive electrostatic field surrounding the bound CO. As a result, the Fe=C=O resonance form dominates and a lower ν_{C-O} of 1940 cm^{-1} is observed (*62*). Structures generated in ChemDraw.

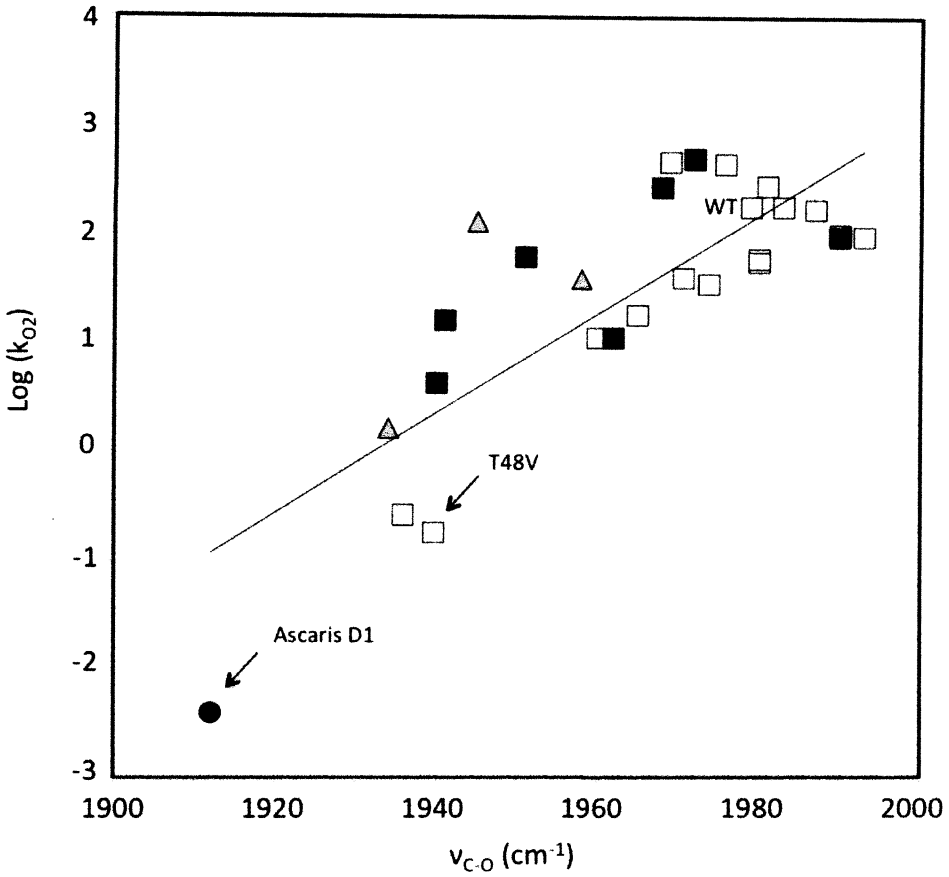


Figure 4.6. Correlation between the C-O stretching frequency (ν_{C-O}) and O_2 dissociation rate constants (k_{O2}) in CerHb. A direct relationship exits between ν_{C-O} and log(k_{O2}) for CerHb mutants (Data from Table 4.2). CerHb single mutants are represented by open squares, CerHb multiple mutants are represented by filled squares, and SwMb mimics of CerHb are represented by filled triangles. *Ascaris suum* Hb D1 is found at 1912 cm^{-1} represented by a filled circle (*94, 95*). The straight line fit has an R-squared value = 0.63886.

the bound CO stretching frequency as a general empirical predictor of relative oxygen dissociation rate constants for all globins, regardless of the active site amino acid composition (*85*).

4.7 Conclusions – Despite their similar O_2 affinities, it is clear from the FTIR and mutagenesis data in this chapter that the electrostatic environments surrounding the bound ligand in wt CerHb and SwMb are markedly different. In wt SwMb, there is a strong positive electrostatic field adjacent to bound ligands as measured by the low ν_{CO} peak at ~1941 cm^{-1}, whereas in wt CerHb there is a negative electrostatic field surrounding the bound CO resulting in a ν_{CO} peak at 1979 cm^{-1}. The new data for our library of CerHb mutants supports the two key relationships that had previously been established for Mb and Hb: (1) a strong inverse correlation between the $\nu_{C\text{-}O}$ peak and the electrostatic field adjacent to bound CO, and (2) a strong direct correlation between $\nu_{C\text{-}O}$ and $\log(k_{O2})$. This comparison provides a clear picture of how CerHb achieved an oxygen affinity similar to that of mammalian Mb when starting from a high affinity protein with two strong hydrogen bonds between TyrB10 and GlnE7 and bound O_2. The addition of ThrE11 to the YQ active site motif causes the TyrB10 hydroxyl to rotate away from the bound ligand due to interactions with the E11 hydroxyl group. This orientation causes the non-bonded electrons on the TyrB10 O atom to destabilize bound O_2, which partially accounts for the high O_2 dissociation rate constant of native CerHb.

However, these electrostatic effects cannot explain why the absolute values of k'_{O2} and k_{O2} are both 10-fold higher than those of wt SwMb. The reason for the higher rates of binding in wt CerHb must be due to a different pathway or kinetic mechanism for ligand

binding than in SwMb. These differences and the importance of understanding ligand migration pathways in globins are explored quantitatively in the next three chapters.

Chapter 5: Entry and Exit in CerHb: Mutations at the E7 and E18 Helical Positions

*Much of the prose, figures, and references that appear in this chapter were taken from Salter *et al.* (2008) (*74*). The use of quotation marks has been omitted from this section to avoid redundancy.

5.1 Introduction - As in Mb and HbA, the shortest route to solvent from the heme Fe is through the E7 gate, which in CerHb, would involve outward and upward movement of the Gln44(E7) side chain away from the heme plane. The proposed alternative pathway in CerHb involves movement in the opposite direction through an apolar tunnel between the E- and H- helices, which is bounded by Val7(B6), Ala82(G8), Leu86(G12), Leu98(H9), and Ala55(E18) (Fig. 5.1). To examine which pathway is applicable in CerHb, I systematically changed the sizes of the key amino acids at the exit/entry point of the E7 gate (Gln44) and entrance to the apolar channel (Ala55). Gln44(E7) was mutated to Ala, Val, Leu, His, Phe, and Trp, and Ala55(E18) was mutated to Val, Leu, Phe, and Trp. The effects of these mutations on overall bimolecular and internal geminate rates of O_2 and CO binding were measured.

The results for the GlnE7 mutations are complex, indicating that this amino acid affects the extent of steric hindrance and electrostatic stabilization of the bound ligand, but, at the same time, these data do not provide any compelling evidence that ligands enter and exit through the shorter E7 gate pathway. In contrast, the results for the AlaE18 mutants indicate strongly that this amino acid forms part of the opening for ligand entry and exit in CerHb. Based on these results, key mutants were sent to Martino Bolognesi, Marco Nardini (University of Milan), and Alessandra Pesce (University of Genova) who crystallized and determined the structures of the Phe55(E18) and Trp55(E18) CerHb

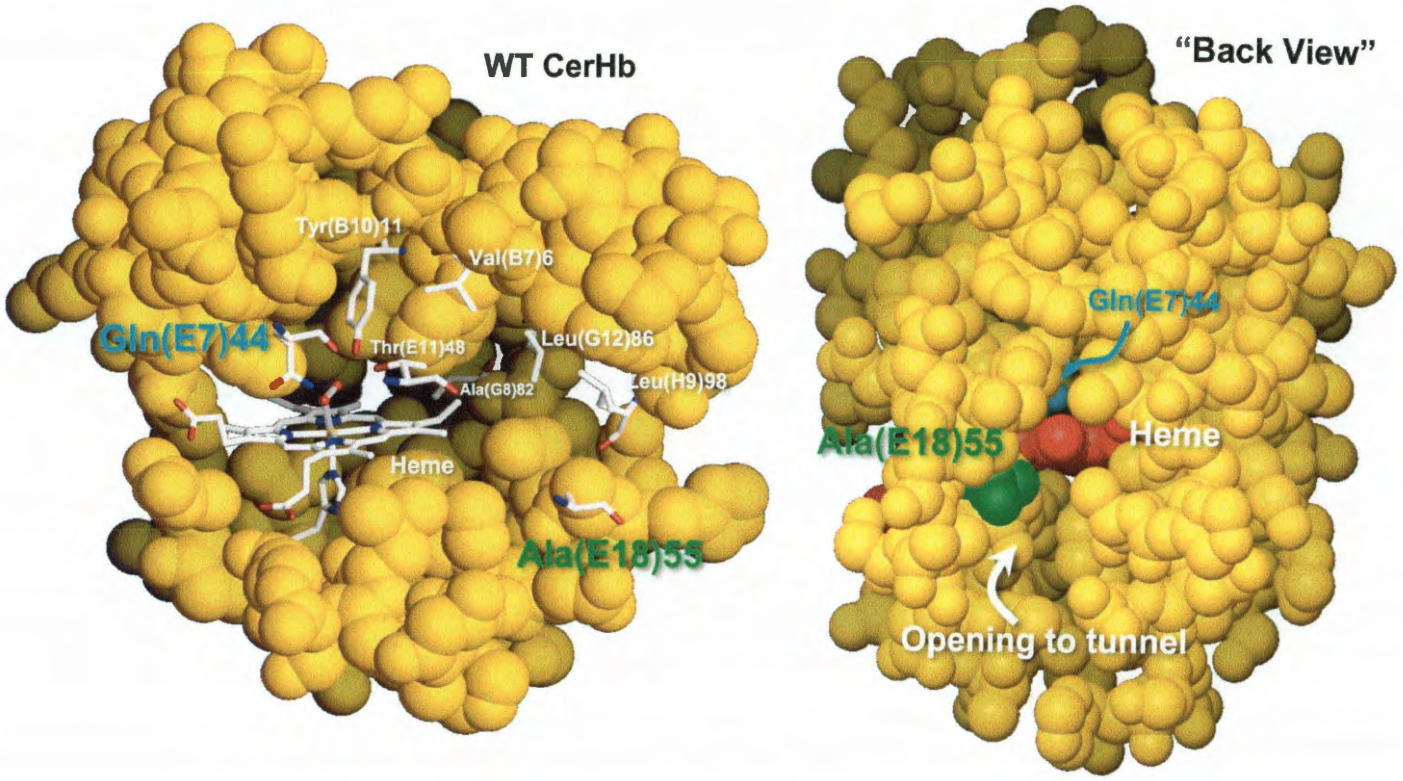


Figure 5.1. Structure of wt CerHbO_2 (PDB: 1kr7) showing side and back views of the molecule. Gln44(E7) is shown in cyan and located directly above the heme propionates. Ala55(E18) is shown in green and located at the end of the E-helix forming part of the open entrance to the apolar tunnel between the E- and H- helices (*74*).

mutants. The structural models showed that these replacements only caused blockage of the entrance to the apolar channel and not marked alterations in the active site. The combined structural and kinetic data suggest strongly that the apolar channel in CerHb is probably the major route for ligand entry and exit. The detailed experimental results supporting this conclusion are provided in this chapter.

5.2 Effects of position 44(E7) CerHb mutants on overall O_2 and CO binding - Bimolecular association (k') and unimolecular dissociation (k) rate constants for ligand binding to Gln44(E7) mutants of CerHb are listed in Table 5.1A. Some of these rate parameters were reported on in previous work performed by Angela Hale as part of her M.A. thesis (*53, 62*). ValE7 and PheE7 mutants were constructed to complete the series of size variants at the E7 gate position and the O_2 and CO binding parameters for all six variants were re-measured to allow a more direct comparison with the same replacements in SwMb, which uses the E7 gate for ligand entry and exit (Table 5.1 and Fig. 5.2).

In general, replacement of HisE7 with smaller apolar amino acids (A,V, L) in SwMb results in dramatic increases in both k'_{O2} and k_{O2} (Table 5.1B). In the case of ligand association, the 5 to 10-fold increases are due to loss of non-covalently bound water (K_{H2O}) and opening of the E7 channel (Fig. 1.4), both of which increase the rate of ligand binding in SwMb. In the case of ligand dissociation, the larger 100 to 1000-fold increases in k_{O2} are due to loss of hydrogen bonding to bound O_2 (Fig. 1.4, $K_{stabilization}$) and greater rates of thermal Fe-O_2 bond breakage (k_{-bond}). The increases in k_{O2} dominate, resulting in large ~ 100-fold decreases in O_2 affinity for the apolar E7 mutants of Mb (K_{O2} values in Table 5.1B and Fig. 5.2C).

Table 5.1. Rate and equilibrium constants for O_2 and CO binding to *A*, position 44(E7) mutants of CerHb and *B*, position 64(E7) mutants of SwMb at pH 7.0, 20 °C. Two phases for O_2 association and dissociation were observed for Trp44 CerHb; roughly 50% of the observed absorbance changes were observed with the fast phase and 50% with the slow phase.

A. CerHb Mutants	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	M (K_{CO}/K_{O2})
Ala44	160	33	4.9	41	0.05, 0.0004	820, 10000	170, 2100
Val44	130	84	1.5	25	0.032	780	510
Leu44	184	96	1.9	36	0.016	2200	1200
Gln44(E7) WT	**230±17**	**190±25**	**1.2±0.2**	**32±4.7**	**0.05±0.03**	**590±100**	**490±110**
His44	85	81	1.0	4.0	0.010	400	380
Phe44	20	170	0.11	1.3	0.054	24	210
Trp44	**10, 170**	**330, 33**	**0.03, 5.2**	5	0.035, 0.011	140, 460	4600, 89

B. SwMb Mutants	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	M (K_{CO}/K_{O2})
Ala64	53[a]	2300	0.02	4	0.061	69	3450
Val64	110	10000	0.01	7	0.048	150	15000
Leu64	98	4100	0.02	26	0.024	1100	55000
His64(E7) WT	**16±3**	**14±3**	**1.1±0.2**	**0.53**	**0.019**	**27**	**25**
Gln64	23	130	0.18	1	0.012	82	455
Phe64	75	10000	0.01	5	0.054	83	8300
Trp64	6.2	87	0.07	1	0.023	28	400

The standard deviations for the wild type CerHb values for k'_{O2}, k_{O2}, and k'_{CO} were obtained from the analysis of 12 completely independent sets of experiments starting with the at least 7 different expressions and purifications of CerHb. The standard deviation for k_{CO} came from the analysis of 16 different determinations. The values of K_{O2} and K_{CO} were calculated from k'_{O2}/k_{O2} and k'_{CO}/k_{CO} and the standard deviations were calculated from standard propagation of error formula (*74*).

[a]The SwMb rate constants were taken from (*89, 91, 93*).

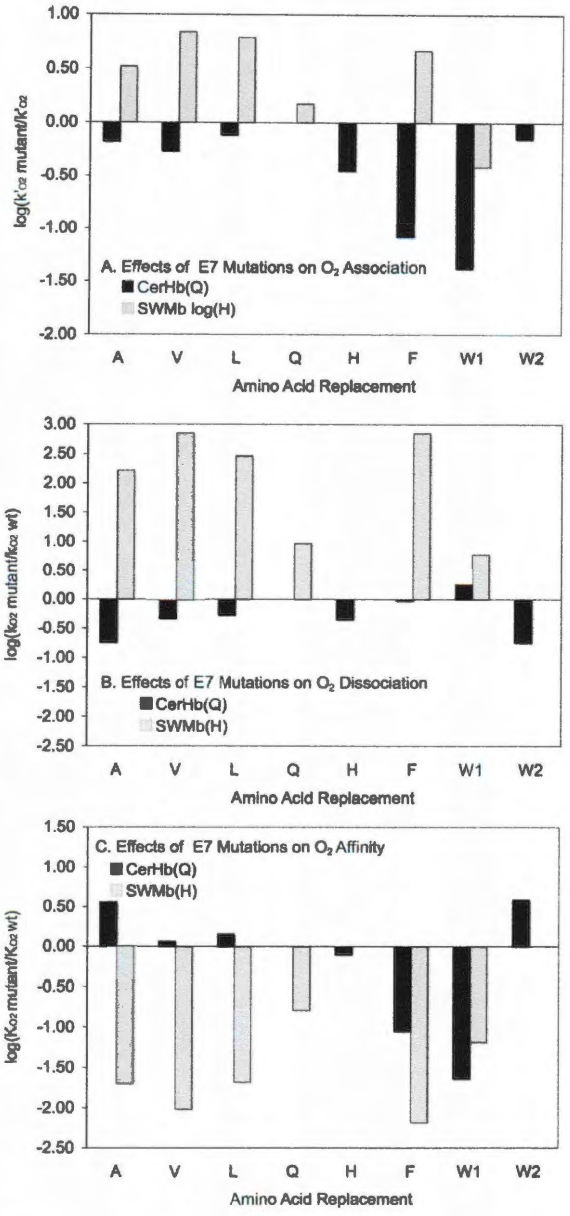


Figure 5.2. Comparison of the effects of E7 mutations on the association, dissociation, and equilibrium constants for O_2 binding to SwMb (*36, 72*) and CerHb. The wt amino acids are His64(E7) and Gln44(E7) in Mb and CerHb, respectively, and the effects are expressed as the logarithm of the ratio of the mutant parameter divided by the wild-type parameter. All rates were measured in 0.1 M phosphate, pH 7, 20 °C. O_2 binding to TrpE7 CerHb showed two phases of roughly equal amplitude. W1 was assigned to the slow component with reduced association and high dissociation rate constants and W2 to the faster reacting component with the low dissociation constant. In contrast SwMb TrpE7 showed only a single slow phase and its values are presented as W1 (*74*).

In contrast to Mb, replacing Gln44(E7) with small apolar amino acids in CerHb results in decreases in the rates of both O_2 association and dissociation and small increases in O_2 affinity (A,V, L black bars in Fig. 5.2). The decreases in k'_{O2} and k_{O2} with decreasing size and polarity at the E7 position argue against this amino acid being part of the pathway for entry and exit because opposite effects, increases in k'_{O2} and k_{O2} would be expected if the E7 channel were enlarged, made apolar, and the gate removed. As shown in Fig. 5.2A, there are significant decreases in the bimolecular rate constants for O_2 binding when the size of the E7 amino acid is increased from Gln to His, Phe, and Trp. However, there are no large concomitant decreases in the O_2 dissociation rate constant (Fig 5.2B), and as a result, there are significant 10 to 3-fold decreases in O_2 affinity for the larger PheE7 and TrpE7 CerHb mutants (Fig 5.2C). These effects for the larger E7 amino acids appear to be due to direct steric and electrostatic interactions with the bound ligand and not necessarily alterations in the speed of entry and escape.

The TrpE7 CerHb mutant shows two distinct and roughly equal phases for O_2 association, O_2 dissociation, and CO dissociation (Table 5.1A). The slower phase for O_2 association shows a rate similar to that for binding to the PheE7 mutant, whereas the faster phase is similar to that of AlaE7 CerHb. The simplest interpretation is that the indole side chain can occupy an inward orientation that directly hinders binding to the Fe atom and an outward orientation that leaves the Fe more accessible, regardless of the pathway for ligand entry and exit. The low k'_{O2} value would correspond to the inward conformer and the high k'_{O2} value would correspond to the outward conformer. Similar effects of the PheE7 and TrpE7 mutants are observed in the α and β subunits of HbA

where inward conformations appear to restrict access to the Fe atom causing marked decreases in both CO and O_2 affinities (*73, 78*).

Using this steric hindrance model, the high rate of O_2 dissociation was assigned to the inward conformer, in which the large indole side chain sterically hinders bound ligands and enhances the rate of Fe-ligand bond breakage. For this model, the calculated O_2 affinity of the inward conformer would be 0.03 μM^{-1}, which is similar to that for the PheE7 CerHb mutant. The outward conformer would have $K_{O2} = \sim 5\ \mu M^{-1}$, which is similar to the affinity of the unhindered AlaE7 mutant. Although compelling, these assignments must be considered tentative in the absence of a crystal structure of the TrpE7 CerHb mutant and further examination. However, Birukou *et al.* (*78*) did determine the crystal structure of TrpE7 deoxy α-chains, and the prediction of an inward conformation was verified for this subunit of HbA.

5.3 Geminate CO recombination in position 44(E7) CerHb mutants - Although the overall kinetic parameters in Table 5.1 indicate that GlnE7 is not acting as a gate for ligand entry and exit, the results are complex and difficult to interpret. To explore the role of this position more completely, time courses for geminate CO recombination in all six E7 CerHb mutants were measured (Fig. 5.3A). In this case, the results are definitive. Neither the rate nor the extent of geminate recombination vary widely with increasing size, from Ala to TrpE7, supporting the idea that the amino acid at the E7 position is not part of the gate for ligand movement into the active site of CerHb.

The largest fraction of geminate recombination is 0.19 (LeuE7) and the smallest is ~ 0.02 (PheE7), with wild-type (GlnE7) being 0.05. The rates of geminate recombination (k_{gem}) do not vary greatly, 60 to ~ 100 μs^{-1}, and, because the extents of geminate

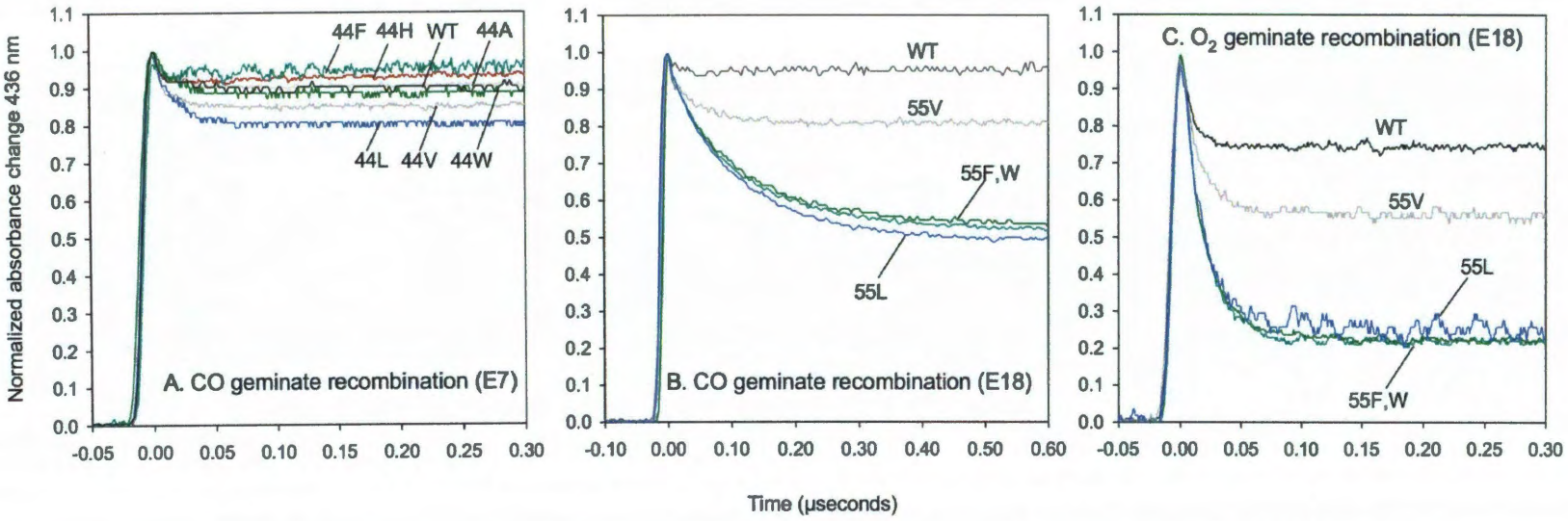


Figure 5.3. Time courses for geminate recombination in mutants of CerHb at 20°C, pH 7.0. *A*, geminate rebinding in position Gln44(E7) mutants of CerHbCO. *B*, CO geminate binding in position Ala55(E18) mutants. *C*, O_2 geminate rebinding in position Ala55(E18) mutants (*74*).

recombination are small (≤ 20%), the values of k_{gem} represent the rates of escape from the CerHb mutants (see Eqs. 5.1 and 5.2 below). Thus, it is clear that changes in the size of the E7 amino acid have little affect on ligand escape. In contrast, Scott *et al.* (*36*) observed systematic decreases in the fraction of O_2 geminate rebinding with increasing size of the E7 position in SwMb, from ~ 0.12 for AlaE7 to 0.72 for TrpE7.

5.4 O_2 and CO binding to position 55(E18) CerHb mutants - As shown in Fig. 5.1, Ala55(E18) is located at the exit/entry point between the apolar channel and solvent between the E- and H- helices. To see if this entry and escape point could be blocked, AlaE18 was replaced with Val, Leu, Phe, and Trp, under the premise that the barrier to ligand uptake would increase with increasing size of the amino acid side chain. As shown in Fig. 5.4 and Table 5.2A, there are strong inverse correlations between amino acid size and the rates of O_2 association and dissociation, with both k'_{O2} and k_{O2} decreasing by > 3 to 4-fold for the AlaE18 to Leu, Phe, Trp mutations and little change in K_{O2}. The concomitant decreases in k'_{O2} and k_{O2} strongly suggest that the mutations are increasing the free energy barrier for entering the protein and only affecting this transition state. This pattern of effects at the E18 position contrasts with the more complex behavior seen when the size of the amino acid at E7 is increased (Fig. 5.4C *versus* Fig. 5.2, respectively).

As shown in Table 5.2, the affinity of CerHb for O_2 shows almost no dependence on the size of the amino acid at the E18 position particularly when considering the uncertainty in the measured value of K_{O2} for wt $CerHbO_2$, which is represented by the standard deviation from ≥ 12 different independent determinations over six years. This observation argues strongly that the mutations only affect movement into the apolar

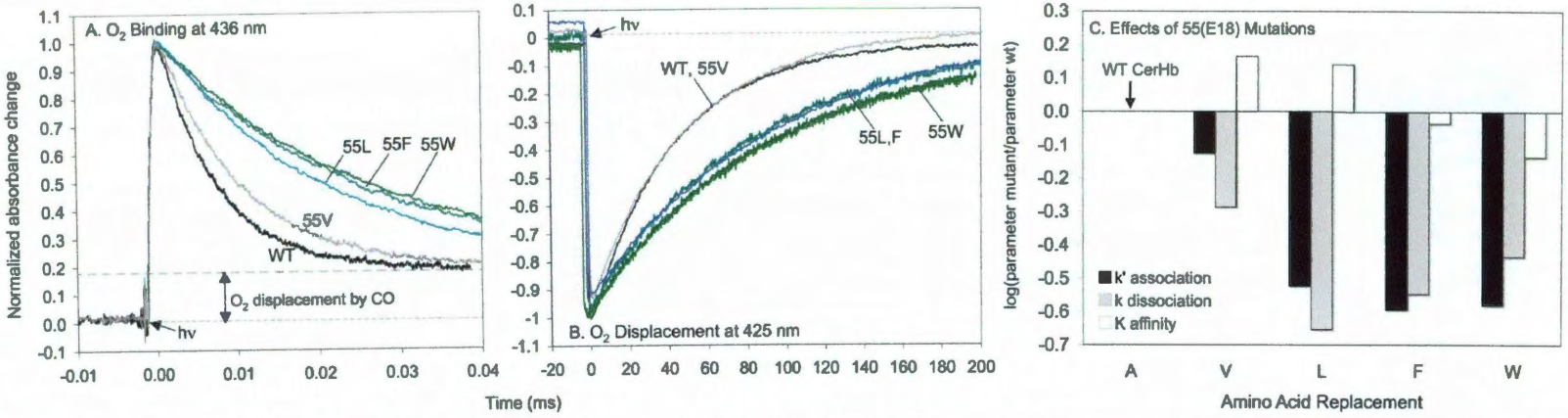


Figure 5.4. Time courses for O_2 binding to wt and 55(E18) mutants of CerHb at pH 7, 20°C. *A*, Bimolecular rebinding measured at 436 nm after photolysis of a mixture of 50 μM CerHbCO at 1250 μM O_2. *B*, Slow O_2 displacement phase measured at 425 nm after photolysis of a mixture of 50 μM CerHbCO, 500 μM CO and 625 μM O_2. *C*, Summary of the effects of the 55(E18) mutations. The expressions used to calculate the association and dissociation rate constants for O_2 binding and release are given in Equations 2.1 and 2.2. The logarithm of the ratio of the mutant parameter to the wild-type parameter are plotted versus amino acid size for the series Ala, Val, Leu, Phe, and Trp (*74*).

Table 5.2. Rate, equilibrium, and geminate constants for *A*, O_2 and *B*, CO binding to position 55(E18) mutants of CerHb at pH 7, 20°C (*74*).

A. O_2 Binding at pH 7.0, 20 °C.

Mutant	Bimolecular binding parameters			Observed Geminate parameters	
	k'_{O2} ($\mu M^{-1}s^{-1}$)	k_{O2} (s^{-1})	K_{O2} (μM^{-1})	k_{gem} (μs^{-1})	F_{gem}[a]
Ala55 WT	230±17	190±25	1.2±0.2	(~130)	0.39
Val55	180	93	1.9	64	0.58
Leu55	72	40	1.8	54	0.83
Phe55	61	51	1.2	55	0.86
Trp55	63	66	0.95	55	0.86

B. CO Binding at pH 7.0, 20 °C.

Mutant	Bimolecular binding parameters			Observed Geminate parameters	
	k'_{CO} ($\mu M^{-1}s^{-1}$)	k_{CO} (s^{-1})	K_{CO} (μM^{-1})	k_{gem} (μs^{-1})	F_{gem}
Ala55 WT	32±4.7	0.054±0.0053	590±100	80±40	0.05±0.03
Val55	41	0.045	910	24	0.16
Leu55	17	0.016	1100	7.6	0.46
Phe55	17	0.029	590	9.3	0.44
Trp55	15	0.029	520	9.4	0.44

[a]The fraction of geminate O_2 rebinding to CerHb was estimated by assuming a "dead-time" of 9 ns for the laser light pulse. This adjustment partially compensates for the amount of geminate rebinding that occurs during the laser pulse, but is only an approximation. The standard deviations for the overall rate constants for wild-type CerHb are described in Table 5.1 and the relative errors ~ ± 10 % for ≥ 12 completely independent measurements are assumed to apply to the parameters for 55(E18) mutants.

tunnel. This interpretation is also supported by the much smaller effects observed for CO association and dissociation. In most globins, CO binding is limited by the internal rate of bond formation or bond breakage and not movement into and out of the protein (*84, 96*). As a result, the overall observed CO association and dissociation rate coefficients are little affected by mutations that only affect entry and exit into the protein as is observed for the position 55(E18) replacements shown in Table 5.2B.

5.5 Effects of position 55(E18) CerHb mutants on geminate recombination - Replacing Ala55(E18) with larger amino acids increases the fraction of geminate recombination, F_{gem}, for both O_2 and CO but slows the rate of this process, k_{gem}, markedly (Figs. 5.3B and C, Table 5.2). In the case of CO geminate rebinding, F_{gem} increases almost 10-fold, from 0.05 to 0.44, whereas k_{gem} decreases 10-fold from ~ 80 to ~ 9 μs^{-1}. Similar increases in F_{gem} and decreases in k_{gem} occur for internal O_2 rebinding, but in this case the values are more difficult to assign since extensive geminate recombination is occurring during excitation and at speeds approaching the rate of decay of the laser pulse (Fig. 5.3C). However, the trends are identical to those seen for CO rebinding. The key observation is that the rate of geminate recombination decreases whereas the fraction of rebinding increases as the tunnel is blocked. These opposite trends can only occur if the rate of ligand escape is decreasing with little change in the rate of internal bond formation as the size of the position 55(E18) amino acid is increased. This interpretation is described quantitatively by Equation 5.1 (Section 5.8)

5.6 Association rate constants for NO binding to the position 55(E18) CerHb mutants - The bimolecular rate of NO binding to globins is determined solely by the rate of entry into the protein because, once captured in the distal pocket, virtually all the

internal ligands bind to the Fe atom before escaping due to the highly reactive free radical nature of nitric oxide. This high reactivity accounts for the large extent of picosecond geminate rebinding of NO in photolysis experiments, the very low overall quantum yield of NO complexes (≤ 0.01), and the large bimolecular association rate coefficients, k'_{NO} (*84*). Scott *et al.* (*36*) have shown that there is a 1:1 correlation between the bimolecular rate of ligand entry, k'_{entry}, calculated from a combined analysis of geminate and overall O_2 rate constants and the values of k'_{NO} determined for over 90 different SwMb mutants. Thus, the rates of NO binding to the E18 mutants of CerHb were measured to examine more directly if the larger amino acids are slowing the rate of ligand entry. As shown in Fig. 5.5 and Table 5.3, the rate of NO binding decreases over 5-fold when Ala55 is replaced with Leu, Phe, and Trp, and the values of k'_{NO} are similar to those for k'_{O2} implying that the rate limiting step for O_2 binding is also movement into the protein and through the tunnel (see also Table 5.3).

5.7 Simple mechanism for ligand binding to CerHb to allow calculation of k'_{entry} - The two-step mechanism shown in Fig. 5.6 was adopted to determine more quantitatively the effects of the E18 mutations on the rates of ligand entry and exit. Deng *et al.* (*53*) have shown that there are multiple geminate states in the distal pocket of CerHb, which interconvert with relatively small enthalpy barriers. These discrete intermediates are easily visualized in FTIR-TDS experiments at cryoscopic temperatures (3-160 K). However, in our experiments at room temperature, simple exponential geminate rebinding is observed, suggesting that only a single intermediate needs to be considered on 10 to 2000 nanosecond time scales.

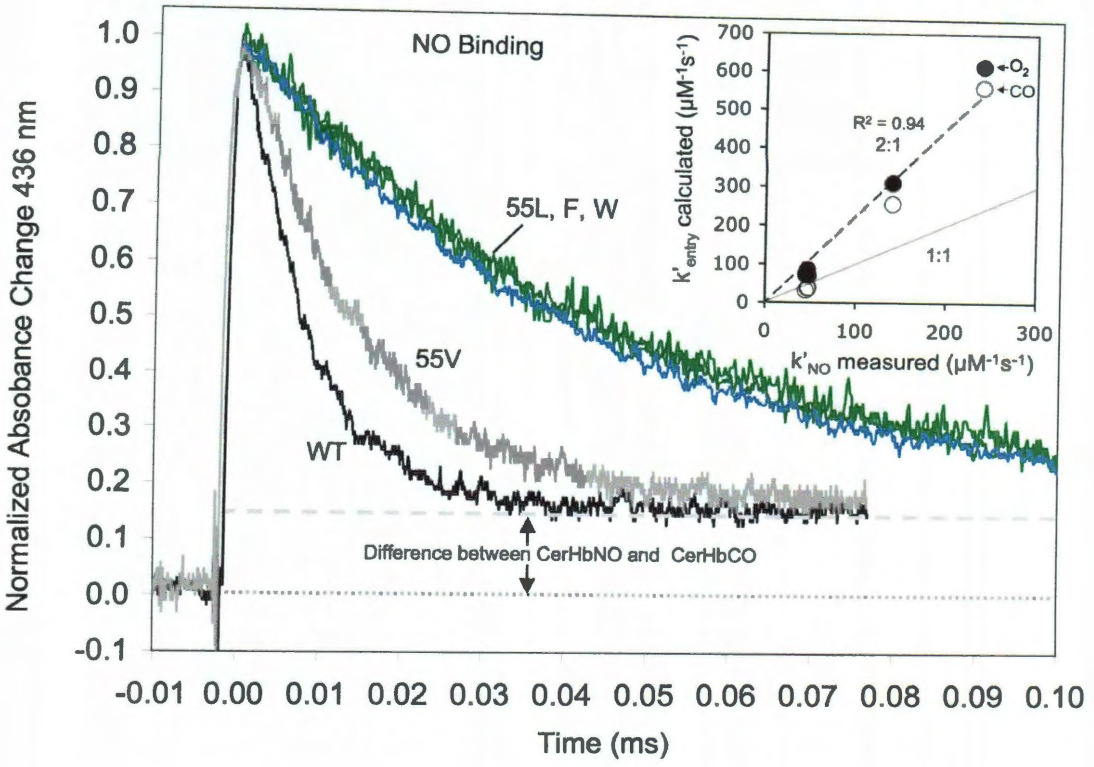


Figure 5.5. NO binding to 55(E18) mutants of CerHb at pH 7, 20 °C. CerHbCO (~50 μM) was rapidly mixing with buffer equilibrated with 1 atm of NO and photolysis was initiated immediately after mixing to generate deoxyCerHb in ~0.5 μs. NO binding was followed at 436 nm. The offset is due to the difference in absorbance between the initial CerHbCO and final CerHbNO complexes. This experiment was repeated with four different NO concentrations and k'_{NO} was determined from the slope of plots of k_{obs} versus [NO]. *Inset*, correlations between the calculated value of k'_{entry} and the measured value of k'_{NO} for wt and the four 55(E18) mutants (*74*).

Table 5.3. Bimolecular rate constants for NO binding and calculated rate parameters for entry, escape and internal bond formation for position 55(E18) mutants of CerHb using the simple two step mechanism shown in Fig. 5.6 and Equations 5.1 and 5.2 (*74*).

A. Kinetic Parameters derived from NO and CO binding at pH 7.0, 20 °C

CerHb	k'_{NO} ($\mu M^{-1}s^{-1}$)	Calculated rate parameters for two step mechanism			
		k'_{entry} ($\mu M^{-1}s^{-1}$)	k_{escape} (μs^{-1})	K_{entry} (μM^{-1})	k_{bond} (μs^{-1})
Ala55(WT)	230±30	640±400	76±40	8	4.0±3.0
Val55	140	260	20	13	3.8
Leu55	47	37	4	9	3.5
Phe55	48	39	5	8	4.1
Trp55	45	34	5	6	4.1

B. Kinetic Parameters derived from O_2 Binding at pH 7.0, 20°C.

CerHb	k'_{NO} ($\mu M^{-1}s^{-1}$)	Calculated rate parameters for two step mechanism			
		k'_{entry} ($\mu M^{-1}s^{-1}$)	k_{escape} (μs^{-1})	K_{entry} (μM^{-1})	k_{bond} (μs^{-1})
Ala55(WT)	230±30	620	79	8	51
Val55	140	310	27	12	37
Leu55	47	87	9	9	45
Phe55	48	71	8	9	47
Trp55	45	73	8	10	47

At ambient temperatures, both geminate and overall ligand binding to CerHb can be analyzed in terms of three reactions: (1) photolysis to the intermediate state, where the ligand is in transient equilibrium between the C, B and tunnel positions, (2) rebinding to the Fe atom from this intermediate state at a rate equal to k_{bond}; and (3) escape from the intermediate state out through the tunnel past the Ala55(E18) position at a rate equal to k_{escape} (Fig. 5.6). The fraction and rate of geminate rebinding, F_{gem} and k_{gem}, are then given by:

$$F_{gem} = \frac{k_{bond}}{k_{bond} + k_{escape}}; k_{gem} = k_{bond} + k_{escape} \qquad \text{Eq. 5.1}$$

$$k_{bond} = k_{gem} F_{gem}; \; k_{escape} = k_{gem}(1 - F_{gem})$$

These fractions and rates of internal recombination are obtained by fitting the observed geminate time courses to a single exponential expression, where the first order rate is k_{gem}, and the normalized amplitude and offset are equal to F_{gem} and $(1\text{-}F_{gem})$ respectively (Fig 2.6 and(*84*)). These parameters can then be used to compute k_{bond} and k_{escape} as shown in Equation 5.1. An expression for the overall bimolecular rate constant for ligand binding, k'_X, where X is the ligand (O_2, CO, NO, *etc*), can be obtained by assuming a steady state approximation (d[I]/dt = 0) for the intermediate on μs and ms time scales. Under these conditions, k'_X is given by:

$$k'_X = k'_{entry} \frac{k_{bond}}{k_{bond} + k_{escape}} = k'_{entry} F_{gem}; \; k'_{entry} = \frac{k'_X}{F_{gem}} \qquad \text{Eq. 5.2}$$

Equations 5.1 and 5.2 and the measured kinetic parameters for O_2 and CO binding in Table 5.2 were used to compute estimates of k'_{entry}, k_{escape}, and k_{bond} for ligand binding to wt CerHb and the four position E18 mutants. The results are shown in Table 5.3 along with the k'_{NO} values. These parameters provide a self-consistent interpretation of both O_2

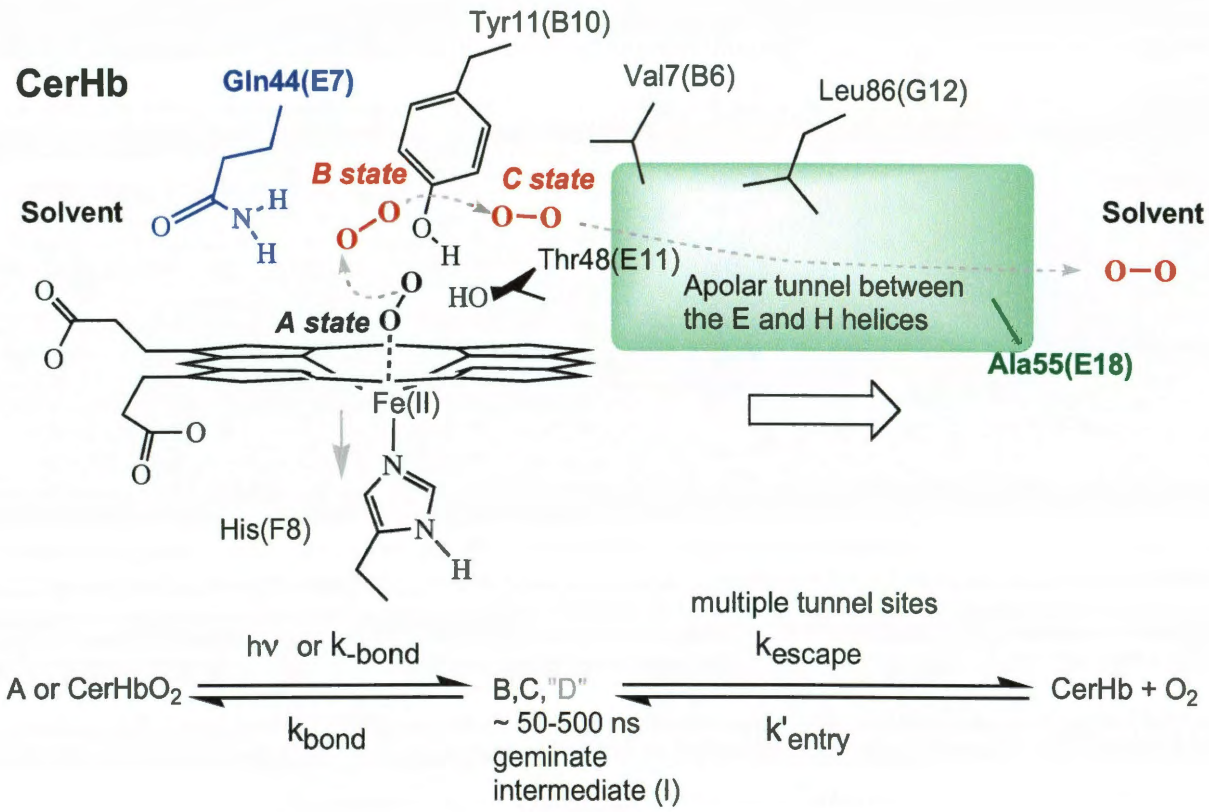


Figure 5.6. Simple two-step mechanism for ligand binding to wild type and 55(E18) mutants of CerHb. Deng *et al.* (*53*) have shown that there are at least three distinct photodissociated states for CO at 3 to 100 K and made tentative assignments to specific cavities within the protein. The B state is close to the iron atom in cavity bounded by the porphyrin ring, Tyr11 and Phe10. The C state appears to be located further into the protein underneath Val7. The D state (not shown) was reported to be in the distal pocket close to Gln44, Thr48, and Tyr11 and thought to be on side path and that ligands either rebind from this state or move into the B to C to apolar pathway for escape (*74*).

and CO binding. The estimated value of k_{bond} for O_2 rebinding is, as expected, roughly 10 times greater than that for CO, ~ 50 μs^{-1} *versus* ~ 4 μs^{-1}. Both values of k_{bond} are independent of changes at the E18 position, demonstrating that the amino acid at this remote location does not influence reaction at the Fe atom. In contrast, the values of k'_{entry} and k_{escape} are roughly independent of ligand, but both rates decrease over 10-fold as the size of the position E18 amino acid is increased (Table 5.3). The correspondence of the absolute values of k'_{entry} and k_{escape} calculated using the O_2 and CO data is remarkably good, considering it is difficult to determine the exact values of k_{gem} and F_{gem} for O_2. For this ligand, significant geminate rebinding is occurring during the laser excitation pulse and the fitted rate is approaching the rate of decay of the pulse light intensity.

As described in the previous section, bimolecular rates of NO binding were measured to verify the computed values of k'_{entry}. NO is expected to show F_{gem} values $\geq$ 0.95, and it was verified that CerHbNO has an extremely small quantum yield for complete photodissociation to solvent (data not shown). Thus k'_{NO} should be approximately equal to k'_{entry}. As shown in Table 5.3 and the inset to Fig. 5.5, there is strong linear correlation between k'_{NO} and k'_{entry}. However, the calculated values of k'_{entry} appear to be ~ 2-fold greater than k'_{NO} for wt and ValE18 CerHb. The cause of this discrepancy is probably the oversimplification inherent in the two-step model (Fig. 5.6). For wt and ValE18 CerHb, the extent of geminate recombination is small; the rate of internal rebinding is large and difficult to define; and there are probably multiple steps that cannot be easily visualized using a 9 ns excitation pulse. In these proteins, the tunnel is open to solvent, and as a result the photodissociated ligands are not easily trapped in the intermediate state. Much more complex expressions define the overall rate constant

when contributions from the rates of interconversion of the B and C states are incorporated into the steady-state expression for k'_{entry}. Evidence for these interior ligand transitions is seen in the kinetic data for mutations in the B and C sites (data not shown), complicating our larger mapping study of the entire pathway (see Chapters 6 and 7). In contrast, blocking the entrance to the apolar tunnel, simplifies the observed results leading to a single, well-defined geminate intermediate. As a result, there is a reasonably close correspondence between k'_{NO} and the calculated values of k'_{entry} for the larger E18 mutants (Table 5.3).

Perhaps the strongest argument in favor of the mechanism in Fig. 5.6 for interpreting the kinetic data is the independence of the estimated equilibrium constant, K_{entry}, for non-covalent capture of ligands in the intermediate state (*i.e.* in the tunnel and the space associated with state B and C) for all the E18 variants. The average computed value is ~ 9.1 ±2.0 M^{-1} for wt and the four E18 CerHb mutants, and does not vary with either ligand or the size of the residue at the entrance to the apolar tunnel (Table 5.3). The lack of dependence on the amino acid at the E18 position fits with the structures of wt, Phe55(E18), and Trp55(E18), all of which show equivalent tunnel volumes of ~ 56 $Å^3$. The value for K_{entry} for ligand capture in wt SwMb is ~ 5.5 ±1.8 M^{-1}, implying that the apolar tunnel in CerHb provides almost twice the volume for non-covalent ligand binding than the active site and Xe cavities in mammalian Mb (*36*). However, in wt SwMb a water molecule is present in the distal pocket reducing the effective volume for non-covalent capture. When this water is removed by replacing His64(E7) with Leu, the value of K_{entry} for SwMb increases to ~10 M^{-1}, and similar values were estimated for ValE7 and PheE7 Mb (*36*). Thus, when the water is removed, the functional volume in Mb, as

measured by the equilibrium constant for non-covalent ligand binding, is similar to that in the CerHb tunnel.

5.8 Crystal structures of Phe55(E18) and Trp55(E18) CerHb - The crystal structures of the Trp55(E18) and Phe55(E18) mutants were determined by Martino Bolognesi, Marco Nardini (University of Milan), and Alessandra Pesce (University of Genova) to visualize whether the large aromatic side chains do physically block the entrance to the apolar tunnel and to examine whether perturbations of the active site occur due to changes at the EF corner of the CerHb tertiary structure. The three-dimensional structures of Phe55(E18) and Trp55(E18) $CerHbO_2$ were refined to final R-factor/R-free values of 18.8%/ 23.0% and 16.7%/ 21.3%, respectively, at 1.8 Å and 1.6 Å resolution, with good overall stereochemistry (for statistics and crystal parameters see Table 4 in Salter *et al.* (*74*)).

The overall tertiary structures of the mutant proteins are unaffected by the mutations at position E18 (Fig. 5.7A). The rmsd calculated between the Cα atoms of the wt and mutated proteins is 0.15 Å for both mutants, and only marginal readjustments are detected at the end of the E-helix and along the adjacent H-helical region (residues 97-104) due to the larger aromatic side chains at position E18. When the two mutant structures are superimposed (overall rmsd of 0.13 Å), the phenyl ring of PheE18 sits approximately in the center of mass of the indole ring of TrpE18 (Fig. 5.7A). The Phe and Trp side chains make 19 and 26 favorable van der Waals contacts, respectively, with adjacent regions of the protein, and the TrpE18 indole Nδ1 atom appears to form a weak hydrogen bond with the Oδ2 atom of Asp104(H15).

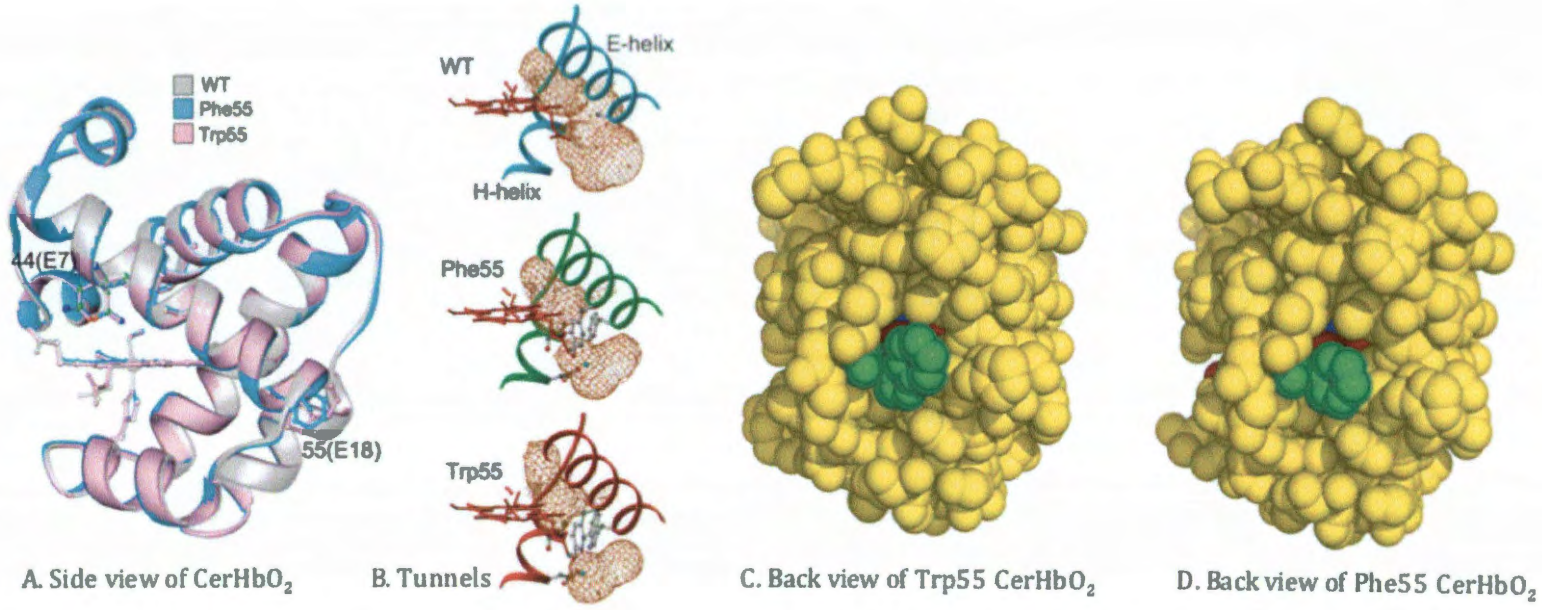


Figure 5.7. Protein matrix cavity in CerHb mutants. *A*, Superimposition of wt (PDB ID: 1kr7, gray), Phe55 (PDB ID: 2vyz, cyan) and Trp55 (PDB ID: 2vyy, purple) crystal structures. Gln44(E7) and 55(E18) side chains are shown in ball-and-stick representation. *B*, Stereo view of the protein core cavities stretching from protein surface to the heme distal site for the wt (cyan ribbon), Ala55(E18)→Phe mutant (green ribbon), and the Ala55(E18)→Trp mutant (red ribbon). Protein cavities have been calculated with a 1.4 Å radius probe using the program SURFNET (*97*), and represented as a net in brown color. Side-chains of the mutated residues at topological position E18 are shown in ball-and-stick representation. For comparison *C*, a back view of $CerHbO_2$ Trp55 is displayed in spheres to demonstrate the blockade of the apolar tunnel by the indole ring designated by green spheres (*74*) and *D*, a back view of $CerHbO_2$ Phe55 where the phenyl ring is displayed in green spheres.

The PheE18 and TrpE18 side-chains fit into a surface cleft, located between the end of the E-helix and the beginning of the H-helix, bordered by surface residues His100(H11), Asp104(H15), and by main chain carbonyl groups of the 55–62 region (EF helical corner, Figs. 5.1 and 5.7). In wt CerHb this cleft is "open" and provides access to an elongated tunnel, lined by the hydrophobic residues Val7(B6), Phe10(B9), Ile52(E15), Ala55(E18), Leu86(G12), Leu98(H9), Ala101(H12), Ile102(H13), Ile105(H16), by hydrophilic residues Tyr11(B10), Thr48(E11), and Tyr51(E14), and by the heme B-methyl and B-vinyl groups (See Figure 7.1B). The surface aperture of the cleft is bordered by residues Ala55(E18), Ala60(F1), Ala62(F2), and Asp104(H15) and has a diameter between 6.8- 7.7 Å (Fig. 5.8). Thus, the major structural effect of the Ala55(E18) to Phe and Trp mutations is blocking access to the tunnel from the solvent phase (Fig. 5.7C).

The blockage caused by the E18 mutations is shown in Fig. 5.7B, where protein cavities are shown as brown nets and were calculated using a 1.4 Å radius probe and the program SURFNET (*97*). In wt CerHb, the apolar tunnel connects directly to the aperture between the E- and H- helices. In contrast, the entrance is almost fully blocked in by the aromatic rings of the PheE18 and TrpE18 side chains. The interior tunnel in the mutants appears as an isolated cavity of roughly the same size (57 $Å^3$) as that found in the wt protein (59 $Å^3$). In both mutants, the tunnel entrance hosts two water molecules hydrogen bonded to Ala101(H12) and Tyr51(E14), and to Asp104(H15) (Fig. 5.8). The presence of these water molecules suggest dynamic opening of the channel, even in the presence of the TrpE18 indole side chain. This interpretation is supported by the observation that rates of O_2 association and dissociation seem to reach a lower limit when the size of the

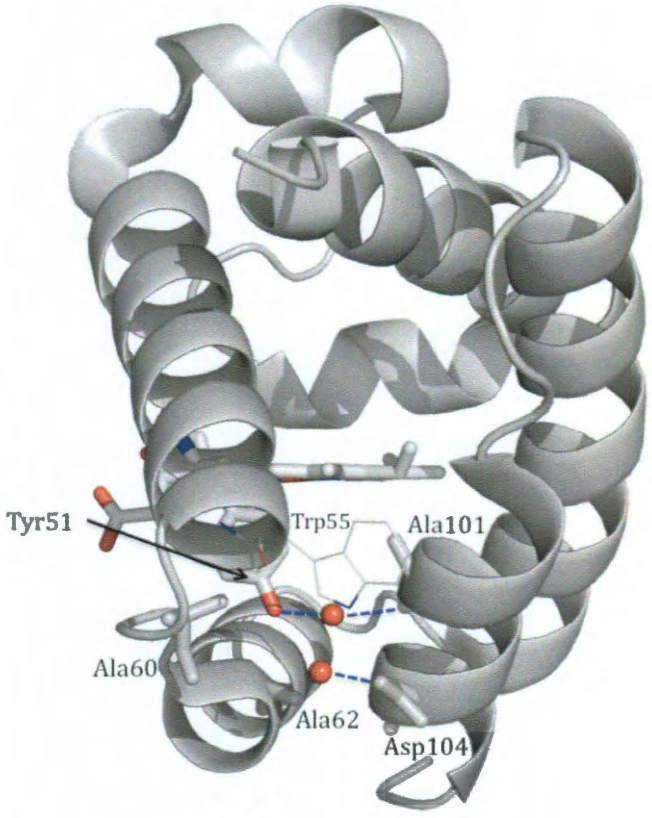


Figure 5.8. Surface aperture of the cleft in Trp55(E18) CerHb (PDB ID: 2vyy). Residues Ala55(E18), Ala60(F1), Ala62(F2), and Asp104(H15) surround the opening of the apolar tunnel in CerHb. Structural representation of the TrpE18 CerHb mutant was chosen to demonstrate the presence of two water molecules (red spheres) hydrogen bonded (blue dashed lines) to residues Ala101, Tyr51, and Asp104 in both Trp55(E18) and Phe55(E18) (PDB ID: 2vyz) mutant models. The carbon backbone of the protein is represented by gray sticks, key amino acid residues are represented by sticks, and the sticks of Trp55 mutation have been reduced in radii so that residues behind this position may be visible from this back perspective. From this perspective, it can be easily imagined that Ala55 reveals an "open" cleft. Protein structure generated in MacPyMol.

side chain at E18 reaches that of Leu, and no significant additional lowering occurs for Phe or Trp (Table 5.2). In addition, the fractions of CO and O_2 geminate recombination do not increase to 1.0 but limit off at values of ~ 0.5 and ~ 0.85, and imply either the presence of alternative routes or a reasonably high frequency of opening the tunnel, evenin the presence of PheE18 and TrpE18 side chains (Table 5.2).

5.9 Multiple conformations of Gln44(E7) - Even though there appear to be few, if any, changes in the positions of the Cα atoms in the N-terminal region of the E-helix and the CD corner (Fig. 5.7A), the amide side chain of Gln44(E7) is found in two conformations (each refined to 0.5 occupancy) in both the PheE18 and TrpE18 mutant structures (Fig. 5.9). The electron density clearly defines two positions of roughly equal occupancy, which will affect electrostatic fields in the active site and could potentially influence rates of O_2 binding and the FTIR spectra of bound CO. This double conformation is particularly well defined in the TrpE18 CerHbO_2 structure, perhaps due to the higher resolution X-ray diffraction data (1.6 Å). The innermost Gln44(E7) conformer exactly matches the rotamer found in the oxygenated wt protein, and appears to form a hydrogen bond between the Gln44(E7) Nε2 atom and the O1 of the bound ligand (O1-Nε2 distances of 2.61 Å and 2.66 Å in the PheE18 and TrpE18 mutants, respectively). The Gln44(E7) Nε2 atom may also weakly interact with the Tyr11(B10) hydroxyl Oζ atom (distances of 3.26 Å and 3.28 Å in PheE18 and TrpE18 mutants, respectively) and with the Thr48(E11) Oγ1 atom (distances of 3.36 Å and 3.25 Å in the PheE18 and TrpE18 mutants, respectively).

In the second outermost conformation, the Gln44(E7) side-chain is rotated outward towards the solvent interface by ~ 45° about Cα-Cβ and the amide NH2 and O

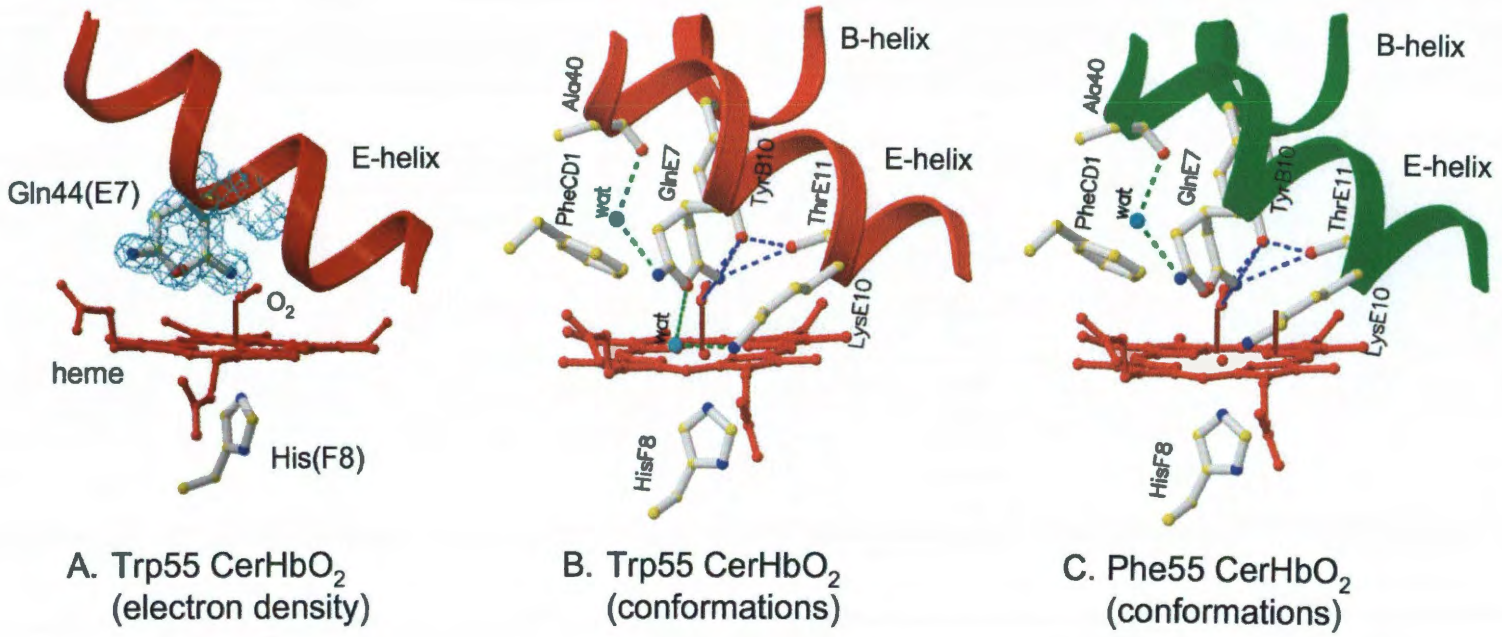


Figure 5.9. The distal pockets of the Trp-55 (*A* and *B*) and the Phe-55 (*C*) mutants. Residues Phe10(B9), Tyr11(B10), Phe25(CD1), Ala40(E3), Gln44(E7) (in double side-chain conformation), Lys 47(E10), and Thr48(E11), the dioxygen molecule (red), the heme group (red), and part of the polypeptide chain path are shown in ball-and-stick and ribbon representations. The proximal HisF8 residue and selected water molecules (see text) are also shown. Hydrogen bonds are drawn as dashed lines (blue for between protein atoms and green for interaction between protein atoms and water molecules). Weak electrostatic interactions may also occur between the innermost Gln44(E7) atom and the Tyr11(B10) hydroxyl but are not shown (*74*).

atoms have been modeled to switch positions by a 180° rotation about Cβ-Cγ (Fig. 5.9B). As a consequence, polar interactions with the bound ligand, Tyr11(B10), and Thr48(E11) are lost. However, in the outer conformation of TrpE18 CerHbO_2, the Gln44(E7) Nε2 atom appears to be part of a hydrogen bonding network that connects the carbonyl of Ala40(E3) to the Nζ of Lys47(E10) through well-defined water molecules, and there may also be electrostatic interactions with the heme propionates. This network is less well defined in the PheE18 structure.

Note that in all the structures, the Thr48(E11) Oγ1 atom is accepting a strong hydrogen bond from the hydroxyl proton of Tyr11(B10) (2.48 Å and 2.56 Å in PheE18 and TrpE18 CerHbO_2, respectively). Thus, bound ligands are experiencing a strong negative field from the non-bonded electrons of the Oζ atom of Tyr11(B10), which accounts for the moderate O_2 affinity (Table 5.2 and Fig. 4.5A) and high stretching frequency of bound CO (~1980 cm^{-1} (*53, 62*)) for wt CerHb and all five position E18 mutants. As described in Chapter 4, mutation of Thr48(E11) to Val causes a dramatic ~ 200 fold increase in O_2 affinity and shift of ν_{C-O} to ~ 1930 cm^{-1} which signifies rotation of the Tyr11(B10) proton back to the bound ligand, creating a strong and favorable positive electrostatic field (Fig. 4.5B). This key role of Thr48(E11) in regulating O_2 binding in CerHb and a complete discussion of the associated electrostatic interactions is given in Chapter 4 and Pesce *et al.* (*62*).

The outward rotation of Gln44(E7) in the E18 mutants is puzzling for two reasons. First, it is not clear how mutations at the end of the E-helix cause this conformational shift of the E7 side chain without causing a significant effect on the position E7 Cα atom. A simple interpretation would be that the larger Phe and TrpE18

side chains sterically clash with the adjacent H-helix, decreasing the angle of the E-F corner, pushing the E- helix slightly toward the distal surface of the heme, and forcing the GlnE7 side chain out of the active site. The inner surfaces of the E- and F- helices are slightly closer in the TrpE18 structure, but the changes are very small and this idea remains speculative. However, as will be discussed in Chapter 6, similar multiple conformations of Gln44(E7) are observed in wt met CerHb and in Leu86(G12)Ala mutants containing bound Xe.

Second, if the E7 residue served as a gate for ligand entry and exit, then the outward movement of the GlnE7 side chain would be expected to increase the rates of O_2 association and dissociation and decrease the fraction of geminate recombination by enhancing the rate of ligand escape. However, the PheE18 and TrpE18 mutations cause the opposite effect, decreasing k'_{O2} and k_{O2} and increasing the fraction of internal rebinding. In addition, the outer Gln44(E7) conformer does not create a significant opening to solvent, even though there is net movement away from the Fe atom (Fig. 5.9). In wt SwMb, protonation of the imidazole side chain of His64(E7) at low pH causes an upward and outward rotation into a truly "open" conformation, creating a direct channel from solvent to the Fe atom (*98*). A similar opening is created by the Phe46(CD4) to Val mutation in MbCO and deoxyMb (*99*). In contrast, access to the Fe atom is still blocked in the outer Gln44(E7) conformer of Trp55(E18) CerHb (Fig. 5.9).

The other remarkable result in Table 5.2 is that the O_2 affinity of all the E18 mutants are, within the error limits (± 20%), virtually the same as that of wt CerHb (Table 5.2), and the stretching frequency of bound CO is unaffected by the PheE18 and TrpE18 substitutions (*74*). These results suggest that the exact conformation of

Gln44(E7) has little effect on either the rate or equilibrium constants for ligand binding. This conclusion is partially supported by the results for the position E7 mutants in Table 5.1. The AlaE7, ValE7, and LeuE7 mutations have little affect on the rate and equilibrium constants for ligand association and, unexpectedly, cause small to moderate decreases in the rate of O_2 dissociation. However, the PheE7 and TrpE7 mutations cause marked decreases in k'_{O2} with little change in k_{O2}. The simplest interpretation of the latter results is that the large aromatic side chains occupy the inner conformation and sterically hinder access to the Fe atom and the bound ligand. Further multiple mutations and structure determinations are needed to verify this interpretation; however, it is clear that the naturally occurring Gln44(E7) side chain does not appear to be part of the main pathway for ligand entry and escape.

5.10 Conclusions - Taken together, the kinetic and structural data for the position E18 mutants suggest strongly that the major pathway for ligand entry and exit in CerHb is the apolar tunnel between the E- and H- helices, which is shown in Figs. 5.1, 5.6, 5.7, and 5.8. Increasing the size of the amino acid at this position increases the free energy barrier to ligand movement from the solvent interface into the apolar channel. Changing the size of the barrier only affects the rates of entry and exit and not the rate of Fe-ligand bond formation, the equilibrium constant for non-covalent ligand capture, nor the overall affinity of CerHb for ligands. The lack of further decreases in either k'_{entry} or k_{escape} as the size of the E18 amino acid is increased from Leu to Phe to Trp suggests that the small, limiting values of these parameters, $\sim 50\ \mu M^{-1}s^{-1}$ and $\sim 6\ \mu s^{-1}$, respectively (Table 5.3) may represent the rates of ligand movement through alternative pathways, including the E7 gate. These rates are remarkably similar to the values estimated for entry into and exit

from wt SwMb which uses the E7 channel where $k'_{entry} \approx 34 \pm 7\ \mu M^{-1}s^{-1}$ and $k_{escape} \approx 6.3 \pm 1\ \mu s^{-1}$ (*36*).

Chapter 6: Defining the Tunnel with Xe Binding and Internal Mutations

*Much of the prose, figures, and references that appear in this chapter were taken from Pesce *et al.* (*100*). The use of quotation marks has been omitted from this section to avoid redundancy.

6.1 Introduction - As highlighted by the previous structural and functional studies (*42, 53, 62, 74*), a wide elongated tunnel, about 10 Å long, >100 $Å^3$ in volume as defined by a 1.4 Å probe, crosses through the CerHb protein matrix. The tunnel is located between the E- and H-helices and directed from the solvent near residues Ala55(E18) and His100(H11) to the heme distal site where it is terminated by Val7(B6), Phe10(B9), and Thr48(E11) (Fig. 5.1). The tunnel is mainly lined by hydrophobic residues, with its inner part characterized by a diameter varying from a minimum of 5.5 Å to a maximum of 6.9 Å, the narrowest segment corresponding to the site occupied by Leu86(G12) (*42*). Diatomic ligands could enter and exit CerHb through the apolar tunnel, whose access from the solvent space is facilitated by the absence of the A-helix, the wide tunnel cross-section, and the short spans of some of the globin helices. As shown in the preceding chapter, obstructing the tunnel entrance by increasing the size of Ala55(E18) to Phe and Trp causes an ~ 4-fold decrease of both k'_{O2} and k_{O2}, an increase in the fraction of geminate CO recombination, F_{gem}, from ~0.05 to 0.45, and a decrease of the rate of geminate CO recombination, k_{gem}, from ~80 μs^{-1} to ~9 μs^{-1}, as ligands are increasingly trapped in the tunnel and inhibited from escaping. Kinetic experiments on Gln44(E7) mutants indicated that a classical E7-gating mechanism for ligand entry and exit, as found in HbA (*73, 78*) and Mb (*101*), is unlikely to be operative in CerHb.

Apolar cavities and tunnel systems have been suggested in selected families of heme-containing proteins, although, relative to CerHb, these are located at slightly

different positions within the overall globin fold topology (*102-104*). In this chapter the structural and kinetic features of CerHb that relate to ligand diffusion within the middle of the protein tunnel system is the point of focus. At Rice University, George Blouin and I obtained direct experimental evidence of access to the heme pocket *via* the internal tunnel by forcing Xe atoms into CerHb at high pressures and then determining the effects on CO binding. Alessandra Pesce and Martino Bolognesi in Italy then determined the crystal structure of a mutant containing bound Xe (*100*). These structural data were complemented by ligand-binding kinetics measured at Rice University for Leu86(G12)Ala, Phe, and Trp mutants, where these substitutions modulate the shape and dimensions of the protein tunnel. In addition, the Bolognesi group determined crystal structures of the Xe-free oxygenated Leu86(G12)Ala mutant and the aquo-met wt CerHb species, which allowed an examination of the tunnel, rates of ligand binding, and changes in orientation of the GlnE7 side chain. Despite small conformational changes at the E7 position, the Gln gate never opened and ligands seem to move primarily through the tunnel (*100*).

6.2 Crystal structures of Ala86(G12) CerHbO$_2$ and its Xe derivative - To examine the accessibility of the internal regions of the apolar tunnel, mutants at position 86(G12) were constructed and sent to our collaborators in Italy for crystallization. Suspensions of crystals of wt CerHbO$_2$ were equilibrated with high pressures of pure Xe gas. Xe has been used previously to identify cavities in a variety of other Mbs and Hbs (*102, 103, 105-109*). Unfortunately, wt CerHbO$_2$ crystals quickly lose diffraction power after exposure to Xe, even at low pressure and short times. This effect could be due to Xe binding at intermolecular crystal contacts, causing disruption of the crystal lattice, or to

Xe binding within the apolar channel, expanding the protein and indirectly disrupting intermolecular contacts.

To solve this problem, the Leu86(G12)Ala $CerHbO_2$ mutant was crystallized and its structure determined (2xkg). In this mutant the constriction near the center of the tunnel is relieved by removal of three side chain carbon atoms at the G12 helical position. When crystals of AlaG12 $CerHbO_2$ were exposed to high pressures of Xe, diffraction persisted after incubation with ~ 10 atm of Xe, and a full data set was successfully collected at 2.3 Å resolution (2xkh, Fig. 6.1). The resulting structure was refined to a final *R*-factor of 19.4 % and *R*-free of 26.6 %, respectively, and compared to the independently determined structure of the Xe-free L86A $CerHbO_2$ mutant, for which a full data-set was collected to 1.6 Å resolution and a model refined to final *R*-factor and *R*-free values of 15.4 % and 18.6 %, respectively (Fig. 6.1, (*100*)).

The backbone structure of Xe-free AlaG12 $CerHbO_2$ is virtually identical to that of wt $CerHbO_2$, with an rms deviation of 0.15 Å calculated for all 109 Cα atom pairs. Other than the replacement of the *iso*-butyl group at position G12 with a methyl carbon, the only significant structural differences between the mutant and wt structures are small rotations of the side chains of Leu98(H9) and Ile101(H13). The net result of these changes and the Ala replacement is an increase in the tunnel inner diameter and volume by ~1.9 Å and 36 %, respectively. This increase either promotes preferential Xe binding in the channel versus binding to external sites that disrupt the crystal lattice or facilitates internal Xe binding without requiring expansion of the protein structure and subsequent alterations in intermolecular packing.

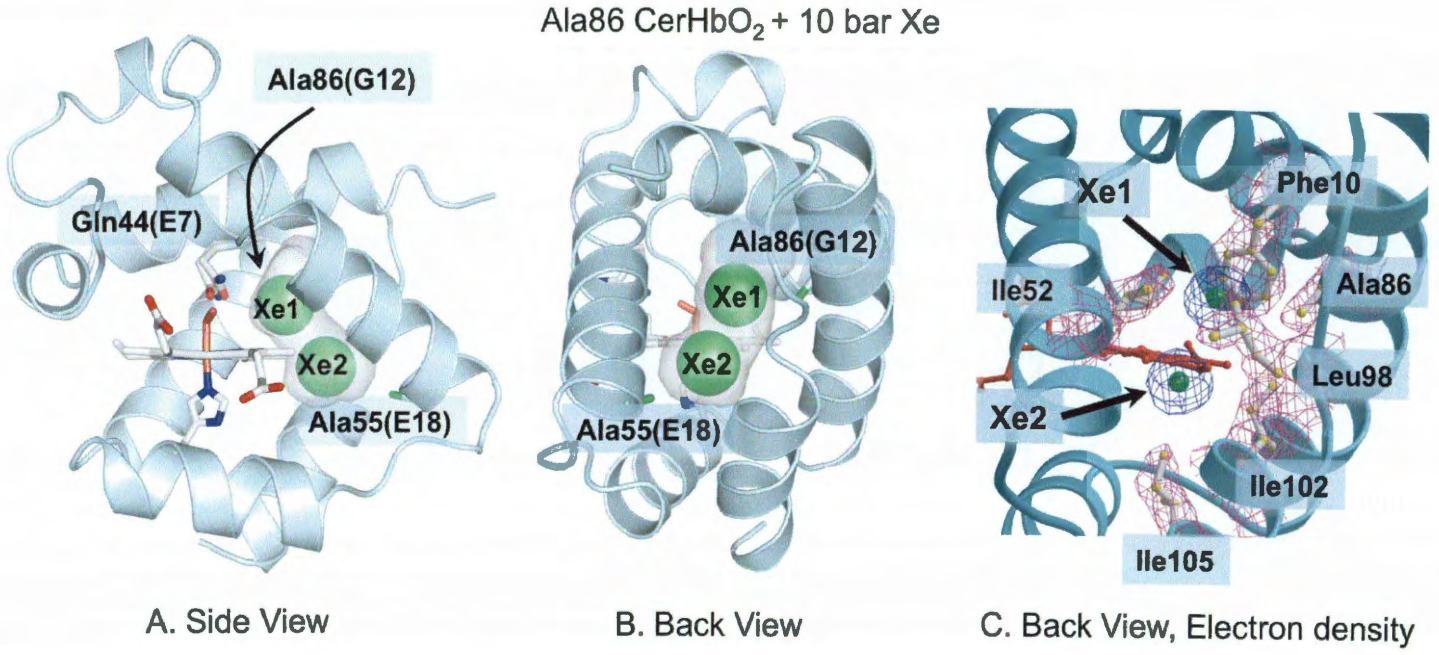


Figure 6.1 Effects of 10 atm Xe on the structure of L86A CerHbO$_2$ (PDB code 2xkh). *A,* Side view looking through the heme propionates into the protein interior. *B,* Back view looking between the E and H helices through the apolar tunnel. The empty space in the tunnel is highlighted in gray. The heme group, proximal His, and Gln44(E7) are shown as sticks with O atoms in red, N atoms in blue, and C atoms in white. The Xe atoms are shown as green spheres at half their van der Waal radii to indicate partial occupancy and designated as Xe1 and Xe2. The Ala86(G12) and Ala55(E18) methyl side chains are shown as green sticks. *C,* Electron density maps of the amino acids and Xe atoms lining the tunnel using the same orientation as in panel *B*. Amino acid side chains are shown as ball and stick models and the Xe atoms drawn as green spheres. The electron densities (2Fo-Fc map contoured at 1s) are shown in magenta for the side chains and in blue for the Xe atoms (*100*).

The heme cavity of Xe-free AlaG12 CerHbO$_2$ is virtually identical to that of wt CerHbO$_2$ (Table 6.1). The general conservation of heme distal site structure and ligand-stabilizing interactions in wt CerHbO$_2$ and in the Xe-free AlaG12 mutant is supported by their similar kinetic parameters for diatomic ligand binding (Tables 6.2-6.4) and ensures that this mutant is a reliable model for studying ligand diffusion through the apolar tunnel. The structure of the AlaG12 mutant containing bound Xe is also very similar to that of wt CerHbO$_2$, with an rms deviation of 0.19 Å for all 109 Cα atom pairs. Inspection of the residual difference electron density, after the initial refinement of the Xe-bound mutant, indicated the presence of two Xe atoms (Xe1 and Xe2) at refined occupancies of ~65 % and temperature factors of 36 Å^2 and 22 Å^2, respectively (Fig. 6.1). The Xe1-site is directly adjacent to Ala86(G12) and surrounded by the ring B vinyl carbon atoms of the heme, the terminal side chain atoms of Phe10(B9) and Ile52(E15). The Xe2 site is adjacent to the Xe1 site, surrounded by the ring B methyl group of the heme, Tyr51(E14), Leu98(H9), Ala101(H12), and Ile102(H13) and is near Ala55(E18) and the exit to solvent. Together both atoms fill the empty space located in the protein interior (transparent gray net in Fig. 6.1) and help define the apolar channel. Both Xe atoms appear to be stabilized by favorable van der Waals contacts with the apolar carbon atoms of Val7(B6), Phe10(B9), Thr48(E11), Tyr51(E14), Ile52(E15), Ala55(E18), Ala86(G12), Leu98(H9), Ala101(H12), and Ile102(H13).

Unexpectedly, Xe binding to the AlaG12 mutant induces small, but readily detectable changes in the position of the Gln44(E7) side chain, which alter the hydrogen bonding lattice near bound O$_2$ and the heme propionates (Fig. 6.2). The model for Xe-

Table 6.1. Distances between polar atoms in the distal pockets of L86A $CerHbO_2$, wild type $CerHbO_2$, and wt aquomet CerHb. The *in* and *out* conformation of the Gln44(E7) side-chain are indicated by the superscript "i" and "o" letters (*100*).

	wt $CerHbO_2$ (1kr7)[a]	Xe-free L86A $CerHbO_2$ (2xkg)	Xe-bound L86A $CerHbO_2$ (2xkh)	wt aquomet CerHb (2xki)
Fe---O2 (Å)	1.94	1.97	1.90	
Fe---wat(Fe^{3+}) (Å)				2.14
Fe-O1-O2 angle (Å)	103	110	-105	
Tyr11(B10) OH---O2 (Å)	2.57	2.47	2.46	
Tyr11(B10) OH---water(Fe^{3+}) (Å)				2.87
Gln44(E7) NE2---O1 (Å)	2.60	2.96		
Gln44(E7) NE2---O2 (Å)			$2.65^{i,o}$	
Gln44(E7) NE2---water(Fe^{3+}) (Å)				2.72^{i}, 7.13^{o}
Tyr11(B10) OH---Thr48(E11) OG1 (Å)	2.59	2.62	2.75	2.68
Tyr11(B10) OH---Gln44(E7) NE2 (Å)	3.24	3.33	3.41^{i}, 7.97^{o}	3.59
Gln44(E7) NE2---Thr48(E11) OG1 (Å)	3.53	3.34	3.48^{i}, 7.50^{o}	3.34
Gln44(E7) NE2—solvent water (Å)			2.72^{o}	2.83^{o}
Solvent water---Ala40(E3) O (Å)			2.59^{o}	2.73^{o}

[a]Structure reported in Pesce *et al.* (*60*).

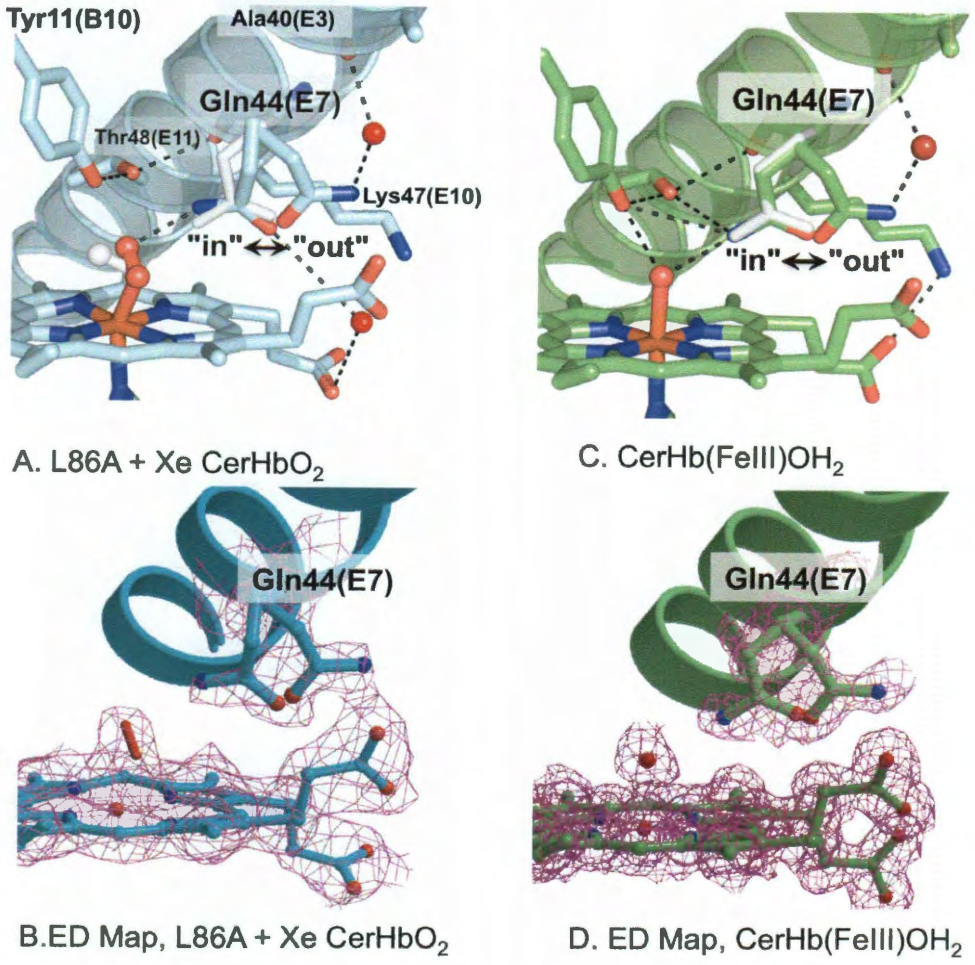


Figure 6.2. Multiple conformations of Gln44(E7) in various structures of CerHb. *A,* Xe-bound L86A $CerHbO_2$ (2xkh, cyan color) showing the distal pocket, E helix, multiple conformations of Gln44(E7), and bound O_2. *B*, The electron density maps, ED (2Fo-Fc map contoured at 1s) for Xe-bound L86A $CerHbO_2$ is shown in magenta for the double conformation of Gln44(E7), the heme group, and the oxygen molecule coordinated to the iron atom for the Gln44(E7) chain refined in two (*in* and *out)* conformations. *C,* wt aquomet CerHb (2xki, green color) showing the heme pocket, E-helix, *in* and *out* conformers of Gln44(E7), and coordinated water. *D*, Electron density maps (2Fo-Fc map contoured at 1s) for wt aquomet CerHb are shown in magenta for the heme group, the heme-coordinated water molecule and for the Gln44(E7) side chain refined in two (*in* and *out*) conformations. In the *upper panels A and C*, the Gln44(E7) side chain conformer from wild-type $CerHbO_2$ (1kr7) is superimposed as white sticks. The position of bound O_2 in wt $CerHbO_2$ is also shown in *panel A* as white sticks. Residues Tyr11(B10), Gln44(E7), Lys47(E10), and Thr48(E11) are represented as sticks; heme-bound O_2 and coordinated H_2O are represented as red sticks; the heme group is displayed as sticks with the C atoms either light blue (panel *A*) or green (panel *C*), the Fe atom orange, and N and O atoms blue and red, respectively; external water molecules are displayed as red spheres; and hydrogen bonds are drawn as dashed black lines (*100*).

bound AlaG12 CerHbO_2 indicates that the second bound O atom rotates outward toward the solvent interface and closer to the terminal O or N atom of Gln44(E7) (Fig. 6.2A). The B factors for the O_2 molecule ($B_{O(1)}$ = 35 Å^2, $B_{O(2)}$ = 38 Å^2) show a higher degree of mobility relative to that in wt CerHb ($B_{O(1)}$ = 23 Å^2, $B_{O(2)}$ = 28 Å^2). This difference in mobility and orientation of bound O_2 is coupled to conformational heterogeneity of the Gln44(E7) side chain.

The innermost Gln44(E7) conformer is close enough to stabilize the ligand by hydrogen bonding if the Nε atom is close to the second bound oxygen atom. The Tyr11(B10) hydroxyl H atom is still hydrogen bonded to the Thr48(E11) hydroxyl O atom, and the non-bonded electrons of the Tyr11(B10) hydroxyl O atom are still close enough to bound O_2 to destabilize it. Thus, the O_2 dissociation rate constant is expected to be large, as is observed for both AlaG12 and wt CerHbO_2. However, in the Xe bound structure of the AlaG12 mutant, Gln44(E7) is too far away to interact electrostatically with Tyr11(B10) and Thr48(E11) in either of its conformations (Fig. 6.2A and Table 6.1).

In the outermost conformation of Gln44(E7) in Xe-bound L86A CerHbO_2, the Nε atom is postulated to move out toward solvent by an ~180 ° rotation about the Cα-Cγ bond. A smaller ~30 ° rotation about the Cα-Cβ bond moves the entire GlnE7 side chain away from the Fe atom. In this conformation, the GlnE7 side chain has lost contact with the ligand and, instead, is stabilized by a hydrogen bonded network linking the Nε atom to the carbonyl oxygen of Ala40(E3) and the heme A propionate carboxyl groups through two well defined water molecules (Table 6.1 and Fig. 6.2A).

Similar *in* and *out* conformations for Gln44(E7) occur in the structure of the wt aquomet CerHb (Fig. 6.2B). The terminal amide side chain atoms in the innermost conformation appear to be interacting with bound water, whereas in the outer conformation the amide side chain has moved away from the coordinated ligand and is again fixed in position by a hydrogen bonding water network involving the carbonyl of Ala40(E3). In both the *in* and *out* conformers of met CerHb, there is an electrostatic interaction between Lys47(E10) and the heme-7-propionate. In the *out* conformers of both Xe bound L86A CerHbO_2 and wt met CerHb, the Gln44(E7) amide side chain still blocks access to the Fe atom by filling the outer portion of the channel between the N-terminal region of the E-helix and the heme propionates. This position contrasts with the truly *out* position of HisE7 in SwMbCO when it is protonated at low pH or when Phe46(CD4) is mutated to Val (*98, 99*).

6.3 Effects of Xe binding on geminate CO recombination - The location of bound Xe in the middle of the apolar channel suggests that escape of photolyzed ligands should be inhibited when samples are equilibrated with high pressures of the gas, and this prediction was confirmed experimentally. As shown in Fig. 6.3, the fraction of geminate recombination (F_{gem}) for both wt and AlaG12 CerHbCO increases almost linearly with increasing Xe pressure, from ~ 0.1 at 0 atm to ~ 0.35 at 15 atm (Table 6.2). Equilibration with 15 atm of N_2 had no effect on F_{gem} for wt CerHbCO, ruling out a non-specific pressure effect (Fig. 6.3A). As observed previously for geminate O_2 rebinding (*36*), there is no effect of 15 atm of Xe on the fraction of CO rebinding in Mb, which remains close to zero (expected value is 0.05, Fig. 6.3B). The linear dependence of F_{gem} on Xe pressure indicates weak binding of the gas to CerHbCO in solution, with a $P_{50} > 15$ atm. The

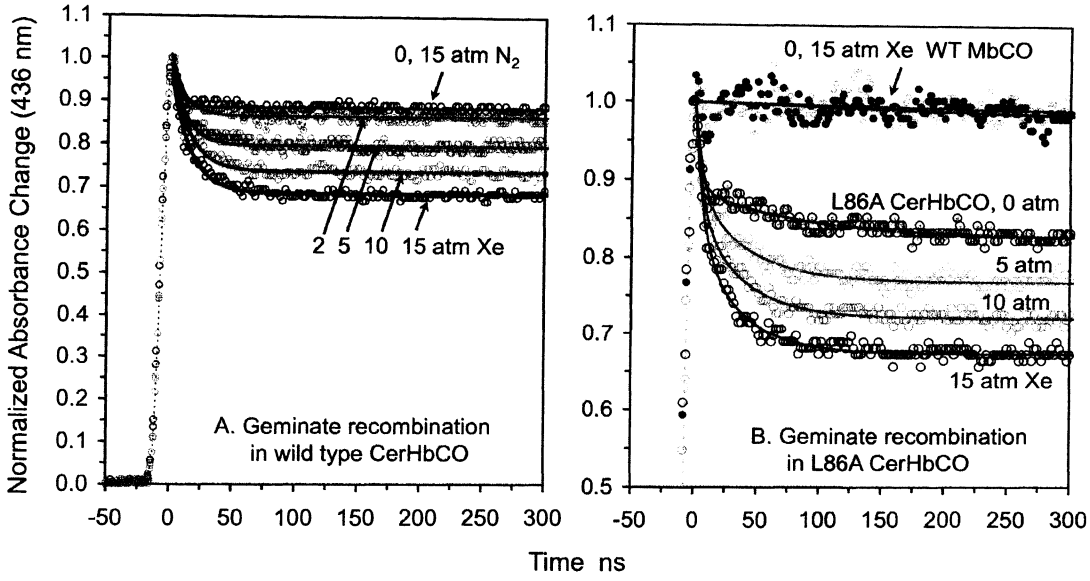


Figure 6.3. Effect of Xe on CO geminate recombination in wild type and L86A CerHbCO at pH 7.0, 20 °C. The smooth solid black lines represent fits to single or double exponential functions with an offset. The gas pressure above each solution is marked next to each time course. There was no effect of 15 atm of Xe on geminate rebinding in wild type sperm whale Mb (WT MbCO, *B*) nor was there an effect of 15 atm of N_2 on geminate recombination in wild-type CerHbCO *(A)* (*100*).

Table 6.2. Rate constants for bimolecular and geminate recombination of CO with wild type and L86A CerHb at pH 7, 20 °C (*100*).

CerHb	k'_{CO} $\mu M^{-1}s^{-1}$	$k_{gem,\ CO}$ μs^{-1}	$F_{gem.\ CO}$	$k'_{entry,\ CO}$ $\mu M^{-1}s^{-1}$
Leu86(wt)[a]	32±5	80±40	0.05±0.03	640±400
Leu86(wt) + Xe (15 atm)	26	63±10	0.29±0.5	90
Ala86	30	≥150, 15	0.13	190
Ala86 + Xe (15 atm)	27	≥150, 28	0.34	80

[a] The numbers for wt CerHb are based on the average of all determinations for the past 6 years (≥ 10 separate preparations, see (5, 8, 9)). The error in k'_{entry} is very large because F_{gem} is close to 0.0, poorly defined, and has an error of ± 60 %. When F_{gem} is ≥ 0.1, the error in k'_{entry} diminishes greatly to ~ ± 20 %.

fractional occupancy estimated from the crystal structure at 10 atm of Xe was ~ 0.65 indicating a somewhat high affinity under the crystallization conditions. Assuming a lower estimate of 15 atm for the P_{50} of both Xe sites, the extent of geminate recombination at full occupancy of sites 1 and 2 is predicted to be ~ 0.7 compared to ~ 0.1 when the channel is completely open. Thus, Xe binding in the channel reduces the extent of CO escape at least 7-fold, strongly supporting the view that the apolar channel is the major pathway for ligand escape.

6.4 Effects of Leu86(G12) CerHb mutants on bimolecular and geminate recombination - The Xe1 and Xe2 sites are located near to side chains at the G12 and E18 helical positions in the apolar tunnel (Fig. 6.1). The effects of increasing the size of the amino acid at the Ala55(E18) position have already been examined and it was found that Leu, Phe, and Trp substitutions trap CO in the tunnel after laser photolysis, increase F_{gem} to ~ 0.50 (Table 5.2B), and decrease the bimolecular rate constant for ligand entry roughly 10-fold from ~500 $\mu M^{-1}s^{-1}$ to ~50 $\mu M^{-1}s^{-1}$ (Table 5.3B). To examine the Xe1 site, ligand binding to AlaG12, PheG12, and TrpG12 mutants was measured (Figs. 6.4 and 6.5 and Tables 6.3 and 6.4).

The PheG12 and TrpG12 CerHbCO mutants show progressive increases in geminate recombination, with F_{gem} for the TrpG12 mutant being roughly the same as that for the TrpE18 variant. However, for the G12 mutants, the rates of geminate recombination remain very large, $\geq$ 100 μs^{-1} compared to that for the TrpE18 mutant, which is 10-fold smaller, ~10 μs^{-1} (Fig. 6.4 and Table 6.5). The simplest interpretation is that TrpG12 keeps the photodissociated ligand closer to the Fe atom, facilitating its

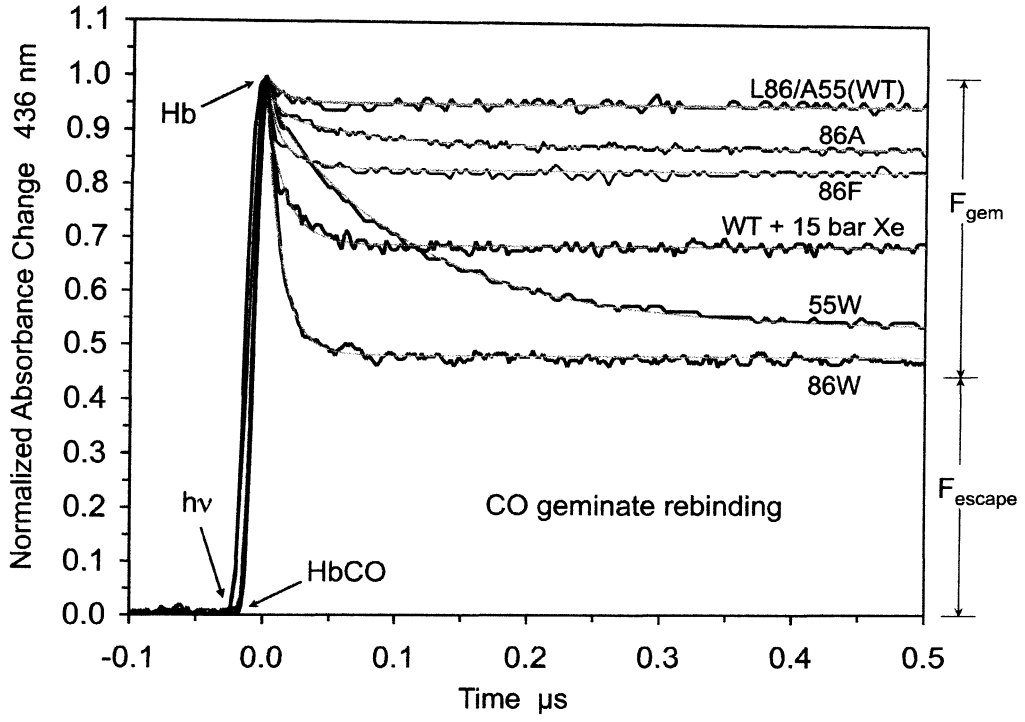


Figure 6.4. Geminate recombination time courses for various CerHbCO mutants at pH 7.0, 20 °C. Geminate recombination time courses for 86A(L86A), 86F(L86F), 86W (L86W), 55W (A55W), and wt CerHbCO in the presence of 15 atm Xe are represented by black lines. The smooth gray lines represent fits to single or double exponential functions with an offset (*100*).

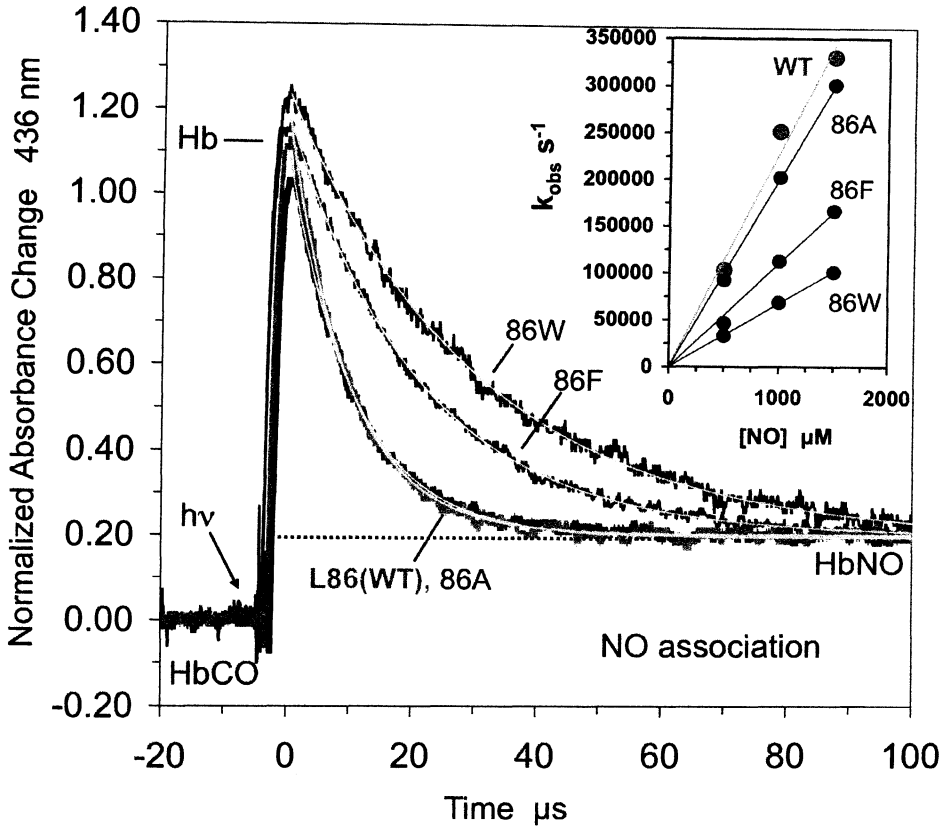


Figure 6.5. NO binding to position 86(G12) mutants of CerHb at pH 7.0, 20 °C. Time courses for NO binding after photolysis of 86A(L86A), 86F(L86F), 86W (L86W), and wt CerHbCO complexes at low [CO] and high [NO] are represented by black lines. The smooth gray lines represent fits to a single exponential function with an offset. The *Inset* shows the linear dependence of k_{obs} (s^{-1}) on [NO] with a *y*-intercept of zero (*100*).

rebinding and lowering the entropic barrier to bond reformation. In the TrpE18 mutant, the entire tunnel is available to dissociated ligands because only the exit to solvent is blocked, and, as a result, both rebinding to the heme Fe and escape from the tunnel are slower.

The rates of O_2 and CO binding to the position G12 mutants are compared to those for wt and TrpE18 CerHb in Table 6.3, and a summary of NO association and calculated bimolecular rates of ligand entry are presented in Fig. 6.5 and Table 6.4. Both the association and dissociation rate constants for O_2 binding decrease 2 to 3-fold when LeuG12 is replaced with Trp, demonstrating that the indole ring limits both diatomic gas escape and entry into the active site, but has little effect on overall affinity. A similar ~ 3-fold decrease in the rate of NO binding is observed, and, as shown in the inset to Fig. 6.5, the observed rates are linearly dependent on ligand concentration.

Estimates of k'_{entry} based on bimolecular association rate constants and F_{gem} values for CO binding to wt and mutant CerHb are given in Table 6.4. (Note $k'_{entry,CO} = k'_{CO}/F_{gem}$, see Equation 5.2.) Equilibrating AlaG12 CerHb with 15 atm of Xe causes a $\geq$ 2-fold decrease in the estimated value of k'_{entry} (Table 6.2). Assuming that the two Xe sites are $\leq$ 50 % occupied, this result suggests that filling the Xe1 and Xe2 pockets would probably decrease k'_{entry} at least 4 to 5-fold. Similarly, the TrpG12 mutation decreases the calculated k'_{entry} value ~ 6 fold compared to the computed value for AlaG12 CerHb (Table 6.4). The magnitude of this decrease is similar to that observed for the TrpE18 mutation, either calculated from CO binding data or estimated from k'_{NO} values. Thus, both geminate recombination and overall ligand binding data, demonstrate

Table 6.3. Rate and equilibrium constants for O_2 and CO binding to position 86(G12) and 55(E18) mutants of CerHb at pH 7.0, 20 °C (*100*).

CerHb	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	K_{CO}/K_{O2}
Ala55/Leu86 (wt)[a]	230±17	190±25	1.2±0.2	32±5	0.054±0.005	590±100	490±110
Ala86	230	153	1.5	30	0.053	570	380
Phe86	120	71	1.7	24	0.024	710	600
Trp86	97	60	1.6	16	0.014	1100	700
Trp55	63	66	0.95	15	0.029	500	520

[a] The numbers for wt CerHb are based on the average of all determinations for the past 6 years (≥ 10 separate preparations, see Table 6.1).

Table 6.4 Observed bimolecular rate constants for NO association and calculated rate parameters for ligand entry into 86(G12) and 55(E18) mutants of CerHb at pH 7.0, 20 °C (*100*).

CerHb	k'_{NO} $\mu M^{-1}s^{-1}$	$k_{gem,\ CO}$ μs^{-1}	$F_{gem.\ CO}$	$k'_{entry,\ CO}$ $\mu M^{-1}s^{-1}$
Ala55/Leu86(wt)[a]	230±30	80±40	0.05±0.03	640±400
Ala86	200	≥150, 15	0.13	190
Phe86	110	≥150, 53	0.17	140
Trp86	67	~80	0.52	31
Trp55	45	9.4	0.44	33

[a] The numbers for wt CerHb are based on the average of all determinations for the past 6 years (≥ 10 separate preparations, see Table 6.1).

unambiguously that the dominant pathway for ligand entry and escape is the apolar tunnel, which is defined by the Xe binding sites.

6.5 The size and accessibility of the apolar channel - The crystal structure of AlaG12 CerHbO_2 equilibrated with ~ 10 atm of Xe demonstrates that the apolar cavity between the E- and H- helices is readily accessible to apolar gases (Fig. 6.1). Xe binding to both wt and AlaG12 CerHbCO causes a significant increase in F_{gem} (Fig. 6.3), demonstrating experimentally that occupancy of the apolar cavity blocks ligand escape. Assuming a partial pressure of 15-20 atm for 50 % saturation of the Xe sites, the predicted extent of CO trapping inside the protein could be as large as 75 % when both Xe sites are filled.

The kinetic results in Fig. 6.3 show that in solution Xe binds within the apolar channel to roughly the same extent in wt and in AlaG12 CerHb, but only the crystals of the AlaG12 variant were stable under high pressures of Xe gas. In general, wt and AlaG12 CerHb have almost identical ligand binding properties and crystal structures, implying that the LeuG12 side chain has only a small effect on the functional properties of the protein. However, wt crystals become disordered in response to Xe binding, presumably because expansion of the channel near the Leu86(G12) position leads to disruption of intermolecular packing.

As in the case of Xe binding, the extent of geminate recombination increases significantly, to ~50 %, when the size of the Leu86(G12) side chain is increased (Fig. 6.4). The PheG12 and TrpG12 mutations also cause significant decreases in both k'_{O2} and k_{O2}, with little change in affinity as would be expected if the pathway for ligand entry and exit were blocked (Table 6.3). The magnitude of this TrpG12 effect is similar to that

observed when exit from the tunnel into solvent is blocked by the TrpE18 mutation (Fig. 6.4 and Table 6.3). In contrast to O_2 binding, there is little change in k'_{CO} because the major rate limiting step for CO binding is internal bond formation with the Fe atom, which is little affected by mutations in the tunnel that are 8 to 12 Å away from the active site. In contrast, the tunnel mutations cause large decreases (4-5 fold) of the rate constant for bimolecular NO binding, which is limited only by movement into the protein (Table 6.4). Similar decreases are obtained for the bimolecular rate constant of ligand entry, k'_{entry}, calculated using k'_{CO}, F_{gem}, and Equation 5.2 (Table 6.4).

6.6 Pathway for ligand binding - The results in this chapter show unambiguously that filling the apolar cavity, with either Xe atoms or large aromatic amino acids at position 86(G12), markedly inhibits ligand entry and exit and increases both the rate and extent of geminate recombination. The simplest interpretation of these results is shown in Fig. 5.6 in the previous chapter, which was first proposed by Deng *et al.* (*53*) and refined by Salter *et al.* (*74*). The apolar tunnel containing the Xe binding sites is assumed to be the major pathway for ligand entry and escape. Laser excitation photodissociates the Fe-ligand bond, generating an initial intermediate with ligand remaining in the distal pocket (state B). From there the ligand can either rebind or move into the apolar tunnel, both of which are rapid processes. However, once inside the tunnel, the ligand can occupy multiple locations in the large space between the distal pocket and the solvent aperture. This larger space slows return to the Fe atom and geminate recombination, allowing most of the ligands to escape into solvent.

When Xe atoms or the indole ring of TrpG12 block the initial portion of the tunnel, ligands are trapped in or very near to the distal pocket and rapidly rebind to the Fe

atom, and large values of k_{gem} and F_{gem} are observed (Fig. 6.4, Tables 6.2 and 6.4). The ligand molecules that move past the bound Xe atoms or the TrpG12 side chain rapidly escape because return to the distal pocket from the exterior portion of the tunnel is blocked. When the tunnel exit is blocked by the TrpE18 mutation, F_{gem} also increases. However, in this case, the rate of internal rebinding decreases markedly because the trapped ligand molecules can access the entire apolar tunnel, making return to the Fe and net escape much slower (Fig. 6.4 and Table 6.4).

This mechanism also explains the results for the AlaG12 mutation. In wt CerHbCO, the LeuG12 side chain does partially sequester ligands close to the distal pocket. Even though this barrier is small, it does separate the two Xe binding sites (Fig. 6.1) and causes the small amount of observed geminate rebinding to be fast in wt CerHbCO. At the same time, the LeuG12 restriction decreases F_{gem} by inhibiting ligand return to the distal pocket from the more exterior portions of the tunnel. When Ala replaces Leu at this position, movement back to the heme distal pocket from the interior of the tunnel occurs more readily. As a result, geminate recombination is biphasic in the AlaG12 mutant. The observed time course shows a small rapid phase with F_{gem} and k_{gem} values = 0.05 and $\geq 100\ \mu s^{-1}$, respectively, which are similar to those for wt CerHbCO, and a second, slow phase with $F_{gem} \approx 0.05$ and $k_{gem} \approx 15\ \mu s^{-1}$. The latter slower phase represents ligand return from the exterior portion of the tunnel, presumably the Xe2 site, at a rate similar to that seen for geminate recombination in the TrpE18 mutant where exit to solvent is blocked.

6.7 Conformational flexibility of Gln44 and the role of the E7 channel - In the crystal structure of wt $CerHbO_2$, the amide side chain of Gln44(E7) appears to be fixed in

position by multiple hydrogen bonding interactions with the bound ligand and the hydroxyl groups of Tyr11(B10) and Thr48(E11). This observation, the existence of the large apolar tunnel between the E- and H- helices, and the effects of mutagenesis led us to propose that ligands enter and exit CerHb through the interior tunnel rather than the more direct E7 gate as in mammalian Mbs and Hbs (Chapter 5). However, when the structure of TrpE18 CerHbO_2 was examined in detail, two conformations for the Gln44(E7) side chain were observed and differed primarily by rotation about the Cβ-Cγ bond. In the *in* conformation, the amide Nε atom donates a hydrogen bond to the bound ligand, whereas in the more *out* orientation, the amide side chain interacts with the carbonyl of Ala40(E3) and the heme propionates through well defined crystallographic water molecules (Fig. 5.7). As shown in Fig. 6.2, two similar Gln44(E7) conformations occur in the Xe bound AlaG12 CerHbO_2 and wt met CerHb structures. Taken together, these structures suggest that the Gln44(E7) side chain is not rigidly held in place but may fluctuate between positions in which the amide side chain is either interacting electrostatically with internal polar side chains and bound ligand when the amide Nε is pointing inward or with Lys47(E10), the carbonyl O atom of Ala40(E3), the heme propionates, and water molecules at the solvent interface when the amide Nε is pointing outward.

When Gln44(E7) is replaced with small amino acid side chains there is no change in the rates of O_2, CO, and NO binding or in the fraction of either O_2 or CO geminate recombination (Chapter 5). The PheE7 and TrpE7 mutations do cause 10-fold decreases in the rates of ligand association but the rates of dissociation stay the same or increase 2-fold. As a result, these mutations cause marked 10-fold decreases in ligand affinity. The

decreases in the association rate constants appear to be due to direct steric hindrance of internal Fe-ligand bond formation by the large aromatic side chains and not to a change in the rate of ligand entry and escape. This conclusion is supported by the lack of an increase in F_{gem} when the size of the amino acid at position 44(E7) is increased progressively from Ala to Trp. F_{gem} remains at 0.05 to 0.1 for all the Gln44(E7) mutants (Fig. 5.3).

In contrast, increasing the size of the amino acids at positions 86(G12) and 55(E18) in the middle and end of the apolar tunnel causes large, progressive, and equal decreases in the rate constants for O_2 association and dissociation. The TrpG12 and TrpE18 mutations and Xe binding within the apolar channel also cause large increases in the fraction of geminate CO recombination (Figs. 6.3-6.5). These effects are unambiguous, indicate that the apolar tunnel is the major pathway for ligand movement into the protein, and demonstrate that the side chains at positions 86(G12) and 55(E18) line the major pathway for ligand entry and exit.

However, the tryptophan tunnel mutations do not completely inhibit ligand binding. Although 4- 5 fold-less than that for wt CerHb, the observed bimolecular rate constants for ligand entry into the TrpG12 and TrpE18 mutants are still large, ~30 to 60 $\mu M^{-1}s^{-1}$ (k'_{NO} and k'_{entry} values, Table 6.5) and similar to the values for both wt and GlnE7 mutants of SwMb (*36*) and the α and β subunits of human HbA (*73*). Thus, 20-25% of the ligand molecules could be entering the active site of CerHb through the E7 channel. The small motions of the Gln44(E7) side chain shown in Fig. 6.2 are very similar to those expected during partial opening of the HisE7 gate in mammalian Mbs and Hbs (*98, 99*). In addition, the barrier heights for ligand entry through the E7 channel

in all three proteins are similar, as reflected in the absolute values of k'_{entry} for native human Hb subunits, native SwMb, and CerHb mutants in which the apolar channel is blocked with aromatic amino acid residues or Xe (*i.e.* 20 to 80 $\mu M^{-1}s^{-1}$).

Chapter 7: Summary: Apolar Tunnels Versus Polar Gates

7.1 Introduction: mutagenesis mapping versus molecular dynamics simulations - In 1990 Elber and Karplus (*110*) carried out locally enhanced simulations of ligand movement in SwMb and observed multiple trajectories with no specific pathway for ligand movement out of the protein. Even the most recent computations still describe multiple pathways in animal Mbs and Hbs with the gated E7 channel always being a minor route (*111-119*). In 1994 Huang and Boxer (*120*) used an elegant random mutagenesis protocol to create and screen a library of ~1,500 single point mutants. Mutations were identified from colonies that produced lysates containing Mbs with kinetic parameters significantly different (~ ±50 %) from those of the wt control. Roughly 10 % of the mutants examined showed differences, and a structural model of Mb was constructed with the positions of the sequenced mutants highlighted. In addition to positions near the heme bound ligand, other regions were highlighted, including interior positions near the Xe cavities and the C-terminal portion of the H-helix, which were interpreted to mean that multiple pathways for CO and O_2 entry and exit do exist, supporting the initial MD simulations (*120*).

Seven years later, building on Huang and Boxer's strategy, Scott *et al.* (*36*) explored the E7 gate and Xe cavity pathways in SwMb by the rational mutagenesis mapping strategy described in chapters 3, 5, and 6, in which small (G or A) and large (F or W) mutations were constructed in ligand accessible cavities, along the major trajectories specified by the MD simulations, and at some of the positions found in the random mutagenesis screen. Each of 90 different mutants was produced in milligram quantities; the kinetics of ligand binding were characterized with pure proteins; and, in

many cases, crystal structures were determined (*32, 91, 99, 121-123*). The mutagenesis mapping results indicated strongly that the major pathway (≥ 75 %) for ligand entry and exit is the short channel that opens by outward movement of HisE7. The discrepancy between the mutagenesis map initiated by Scott *et al.* (*36*) and that of Huang and Boxer (*120*) is not as great as it first appears. The largest changes in their random mutagenesis map are observed at or near the E7 gate, and most of the effects observed at remote positions can be rationalized as false positives, due to mutations that destabilize Mb leading to lysates containing partially unfolded protein. The discrepancy between the calculated multiple trajectories for ligand migration obtained in molecular dynamics simulations and the dominant E7 gate pathway determined experimentally in SwMb is harder to reconcile and implies that internal pathways cannot be "found" with a rational mutagenesis mapping strategy. Elber (*39*) has summarized this dichotomy in a recent review on computational methods for finding ligand trajectories.

As described in Chapters 5 and 6 (*74, 100*), it has been shown that the mini-hemoglobin from the sea worm *Cerebratulus lacteus* has an apolar tunnel leading from the distal pocket to solvent, with an exit point between the E- and H-helices created by the loss of the N-terminal A-helix, and our initial mutagenesis and Xe binding studies suggest strongly that this internal tunnel is the major ligand pathway. However, Orlowski and Nowak (*117*) carried out simulations with CerHb and, as with Mb, obtained multiple trajectories with the apolar tunnel being only one of them. To resolve this discrepancy and determine whether the experimental mutagenesis mapping strategy used for SwMb by Scott *et al.* (*36*) can identify internal pathways, 120 different mutants of CerHb were constructed and the O_2, CO, and NO ligand binding properties were examined. Rates of

ligand entry and exit for this library of CerHb variants were compared to the results for an expanded library of SwMb mutants, extending the initial results of Scott *et al.* (*36*) to regions near the ends of the E- and H-helices. The results of these comparisons are provided in this chapter, which summarizes our views on ligand pathways in globins in general, using SwMb and CerHb as model systems, and on the validity of the mutagenesis mapping strategy to find them.

7.2 Mechanisms for ligand binding to SwMb and CerHb and estimations of k'_{entry} and k_{escape} - A comparison of the distal pockets of CerHb and SwMb is shown in Fig. 7.1. Key amino acid side chains near the E7 gate (Gln44 in CerHb, His64 in Mb), at the Xe binding sites, and along the apolar tunnel in CerHb between the E- and H-helices are shown as sticks and labeled. From these side views, it is clear that the equivalent apolar tunnel in SwMb is blocked both by the A-helix and by larger amino acid side chains located along the interiors of the E-, G-, and H-helices. To define the actual pathway in CerHb, a series of replacements with large (F or W) to small (A) amino acids at positions in or along the E7 channel (Chapter 5), the apolar tunnel (Chapters 5 and 6), and other regions reported to be important based on previous simulations and experimental work with recombinant Mbs and Hbs has been constructed. The overall rate constants for O_2, CO, and NO binding were measured using rapid mixing, laser photolysis methods with O_2/CO mixtures, and flow-flash techniques as described in Chapter 2 (Section 2.5). Internal CO recombination within each mutant was also measured using a 7 ns YAG laser, and the rate of geminate recombination, k_{gem}, the fraction of recombination, F_{gem}, and the overall bimolecular rate constant for CO binding

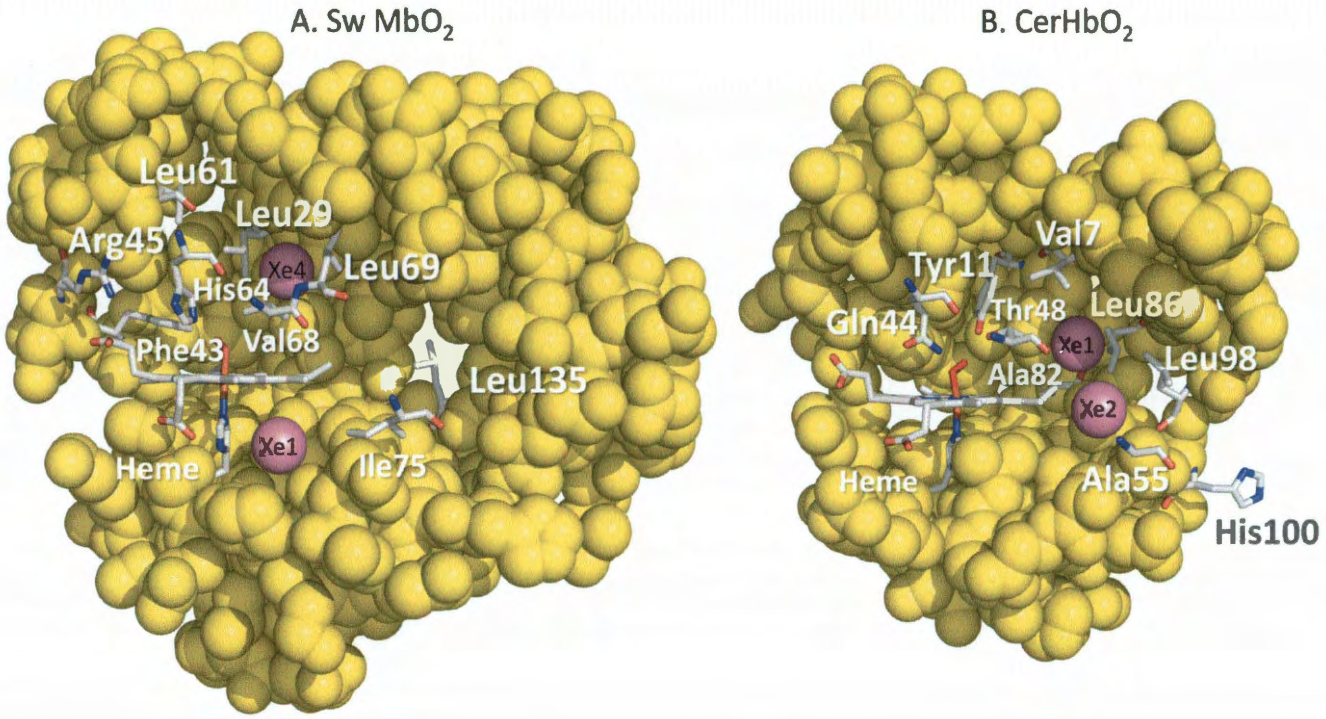


Figure 7.1. Structural representation of Xe-bound WT $CerHbO_2$ and SwMb showing side view of the molecules. In both globins the molecules are oriented from the side with the heme propionates facing left and portions of the protein matrix slabbed away so an interior view is clear. Xe atoms are depicted as violet spheres; key amino acids and the heme prosthetic group are indicated by gray sticks with oxygen in red and nitrogen in blue; the remainder of the protein matrices are represented as yellow spheres. *A*, The two Xe atoms in CerHb are oriented as in the Leu86Ala structure (PDB ID: 2xkh). Key distal pocket and apolar tunnel residues include Val7(B6), Tyr11(B10), Gln44(E7), Thr48(E11), Ala55(E18), Ala82(G8), Leu86(G12), Leu98(H9), His100(H11). *B*, Xe1 and Xe4 of SwMb are included in this depiction to indicate proposed ligand cavities within the protein matrix. Key distal pocket residues include Leu29(B10), His64(E7), Val68(E11), and Phe43(CD1). Additional mutants constructed in this study include mutations made at residues Arg45(CD3), Leu61(E4), Leu69(E12), Ile75(E18), and Leu135(H12). Protein structures generated with MacPyMol.

were used as described below to estimate the first order rate constant for ligand escape and the bimolecular rate for ligand re-entry from solvent. All of these kinetic parameters are listed in Appendix A (Tables A.1-A.6).

In Scott *et al.* (*36*) bimolecular and geminate O_2 binding to SwMb were analyzed in terms of the three-step, side path model shown in Fig. 7.2. The rate constants for intramolecular rebinding, escape, and movement through the protein between sites B and C were obtained by fitting sets of time courses to this model using numerical integration algorithms described in Scott and Gibson (*79*). The geminate rebinding time courses for wt and many of the mutant Mbs can also be fitted well by the sum of two exponential processes and an offset reflecting the amount of escape to solvent (Fig. 2.6). The side path scheme in Fig. 7.2 was chosen because both Xe binding and mutations that fill the Xe1 and Xe4 cavities collapse the time courses to a single rapid phase with little or no change in the total fraction of geminate rebinding (*36, 79*). Similar results were obtained for geminate ligand rebinding to wt *Scapharca inaequivalvis* HbI, where again Xe binding and Phe and Trp replacements in internal cavities caused little change in the total fraction of geminate recombination (Chapter 3 and (*81*)). In the case of HbA subunits, single geminate phases for CO rebinding are observed at room temperature on ns time scales and indicate that there are no side path C states, and the E7 channel also appears to be the major path (> 90%) for ligand entry and escape (*73, 78*). Thus, the side path scheme appears to apply to most vertebrate hemoglobins with a tertiary and distal pocket structure similar to that of SwMb.

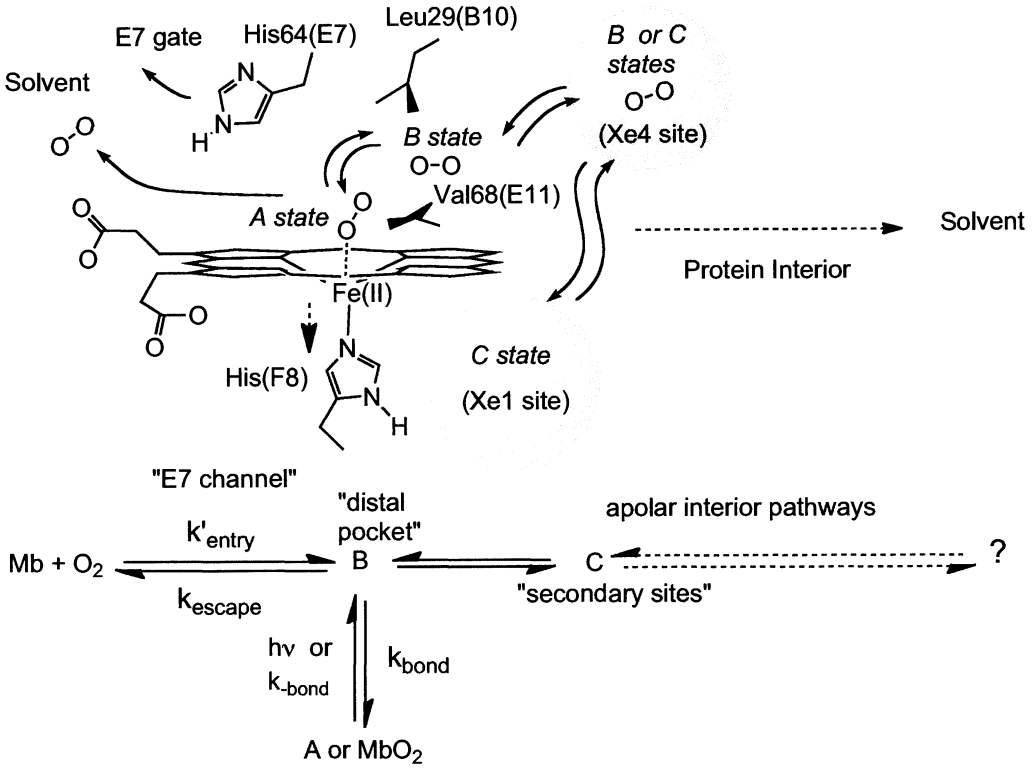


Figure 7.2. Side path ligand binding scheme for SwMb. Ligand enters from the solvent (left) into the distal pocket where it temporarily resides until bond formation (A state) and stabilization occur. Following photolysis, ligand resides in the geminate state (B state) and/or travels to a Xe pocket (C or D state). Ligand must first return to the geminate state from the C and D states in order to escape the protein via the HisE7 gate. Figure taken from (*36*). The figure was generated in ChemDraw.

It is important to note that the experimental values of F_{gem} are model independent and depend only on an accurate definition of absorbance change for the total amount of geminate recombination. Thus, Equations 5.1 and 5.2 can also be used empirically to estimate the rate of entry into any globin, including CerHb, where, in this case, the two-step binding scheme involves a linear sequential set of ligand movements (Fig. 5.6). As discussed by Scott *et al.* (*36*), the steady state expression for k'_X using a linear scheme is still $k'_X = k'_{entry} F_{gem}$, where k'_{entry} represents the bimolecular rate of entry into the first discrete site in the protein, which, in the case of CerHb, would be the apolar tunnel. However, if the barriers along the tunnel and from the tunnel into the distal pocket are small compared to bond formation, the mechanism for CerHb reduces to a two-step process with a single geminate intermediate. Under these conditions, F_{gem} can be approximated as $k_{bond}/(k_{bond} + k_{escape})$.

In the case of CO binding to CerHb, there is evidence of a barrier between the tunnel and the distal pocket implying a linear three-step scheme. However, this extra complexity is only evident at room temperature when restrictions in the tunnel are completely removed as is observed in the Leu86(G12)Ala mutation (*100*) or at low temperatures (*53*). When movement into and through the tunnel is restricted by mutagenesis, the geminate time courses become single exponentials and are readily analyzed as a simple two-step scheme where Equation 5.1 describes the geminate parameters. Expressions in Equations 5.1 and 5.2 were used to estimate values of k'_{entry}, k_{escape}, and k_{bond} for CO binding to all of the mutants of CerHb listed in Tables A.1- A.6 (see Appendix A). These estimations were felt, although empirical, to be the simplest and

closest to the observed experimental data, making the maps shown in Fig. 7.4 (see Section 7.4) experimentally defined and not model dependent.

7.3 Validity of using k'_{entry} and k_{escape} - In principle, k'_{entry} and k_{escape} values should be used to map the ligand pathways quantitatively because these parameters factor out the differences in reactivity (k_{bond} values) between the O_2, CO, and NO. The values of these rate constants should be the same for all three diatomic gases because of their similar sizes and diffusion constants at room temperature. In the case of CO binding to wt CerHb, the relative errors in these calculated parameters are very large because the extent of CO geminate recombination is very small ≤ 0.05 and has a absolute error of $\sim \pm 0.03$ which is roughly the same for all values of F_{gem} measured. To reassure ourselves that the calculated k'_{entry} values are valid for the CerHb mutants, the correlations between k'_{entry}, k'_{NO}, and k'_{O2} were examined. As shown in Fig. 7.3, there is a linear correlation between k'_{entry}, calculated using k'_{CO} and $F_{gem,CO}$, and the experimentally measured value of k'_{NO} for a series of Val7(B6), Gln44(E7), Thr48(E11), Ala55(E18), and Leu86(G12) mutants (Table A.2, Appendix A). Because the fraction of geminate NO recombination is ≥ 0.99 due to its high reactivity with Fe, the rate limiting step for its binding is the bimolecular rate of NO entry into the protein, which should equal the calculated value of k'_{entry}. As shown, the error in k'_{entry} is very large for those mutants with large k'_{CO} and very small $F_{gem,CO}$ values, making it hard to determine the slope of the correlation between k'_{entry} and k'_{CO}. However, even with these errors, there does appear to be a systematic overestimation of k'_{entry} for the CerHb variants with $F_{gem,CO} < 0.1$ (Fig. 7.3A). However, the correlation is strong for mutants where F_{gem} is larger when the tunnel is partially blocked. Similar linear correlations between k'_{entry} and k'_{NO} were observed for all of the

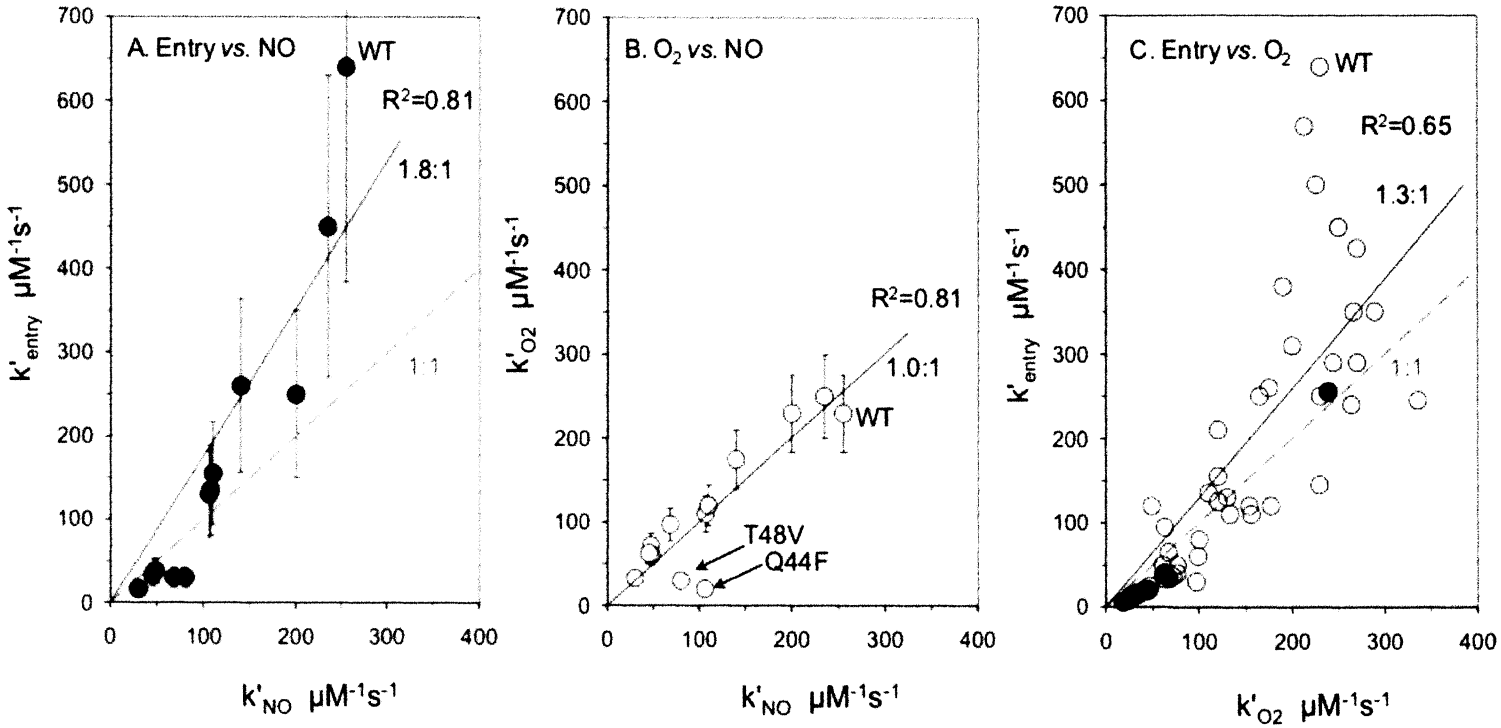


Figure 7.3. Correlation of calculated rates of ligand entry, k'_{NO}, and k'_{O2} in CerHb. *A*, k'_{entry} vs. k'_{NO}, *B*, k'_{O2} vs. k'_{NO}, and *C*, k'_{entry} vs. k'_{O2} for positions Val7(B6)A, F, W, Thr48(E11)V, Ala55(E18)V, L, F, W, Leu86(G12)A, F, W, and Gln44(E7)F.

SwMb mutants reported in Scott *et al.* (*36*), verifying the k'_{entry} values for myoglobin.

As shown in Fig. 7.3B, there is a 1:1 correlation between k'_{NO} and k'_{O2} for the single mutants located in the tunnel implying that O_2 binding is also primarily limited by the bimolecular rate of entry into the protein. The only exceptions are the Thr48(E11)Val mutation, which releases the TyrB10 to sterically inhibit bond formation, and the Gln44(E7)Phe mutation, where the larger PheE7 side chain directly inhibits O_2 binding at the Fe atom. For these mutants, overall O_2 binding is slowed because bond formation is markedly inhibited and k_{bond} becomes similar to k_{escape} causing a decrease in k'_{O2} due to a decrease in F_{gem} and not the rate of entry (Equations 5.1 and 5.2). In the case of NO, its high reactivity with the Fe atom is still much greater than that for ligand escape, keeping ligand entry as the rate-limiting step even for the Thr48(E11)Val and Gln44(E7)Phe mutants.

Fig. 7.3C shows the correspondence between the calculated ligand entry rate using overall and geminate CO binding data with the bimolecular rate constant for O_2 binding. In this case, there is a 1:1 correspondence with all the mutants examined even though the scatter is fairly large. These results, the data in Fig. 7.3A, and correlations between NO binding rates and k'_{entry} values for the SwMb mutants (*36*) support the use of calculated k'_{entry} and k_{escape} values to map ligand pathways and demonstrate that similar conclusions would be obtained if k'_{NO} or k'_{O2} values were used in the mapping analysis.

7.4 Mutagenesis mapping strategy - The underlying premise of our mutagenesis approach is that larger and smaller amino acid side chains should impose larger or smaller kinetic barriers, respectively, for ligand entry and escape if they are located along the primary pathway for ligand binding (Chapters 3, 5, and 6; (*36, 74*)). Thus, the overall

bimolecular rate constant for ligand entry into a mutant with a small amino acid, $k'_{entry,small}$, should be significantly larger than the entry rate constant for a mutant with a large amino acid, $k'_{entry,large}$, at the same on-pathway position. Similar relationships should occur for the rate constants for ligand escape from the same mutants. To quantify these effects, overall effects of large to small mutations on these rate enhancements were computed as (*36*):

$$R_{enhance} = \log\left(\frac{k'_{entry,small}}{k'_{entry,large}}\right) + \log\left(\frac{k_{escape,small}}{k_{escape,large}}\right) \qquad \text{Eq. 7.1}$$

Large to small mutations at positions on the major pathway increase both k'_{entry} and k_{escape} leading to positive values for both logarithmic terms in Equation 7.1, which, when added together, generate large $R_{enhance}$ values. In contrast, mutations at positions distant from the ligand trajectory have little or no effect. Mutations near the bound ligand often have large effects on overall ligand binding parameters due to either direct steric hindrance or electrostatic stabilization of the bound ligand, but in many cases, there are only small or compensating effects on entry and escape.

To complement some of the new helical positions examined in CerHb, I constructed and characterized eight additional SwMb mutants: Arg45(CD3)Trp, Leu61(E4)Trp, Leu69(E12)Ala, Leu69(E12)Trp, Ile75(E18)Ala, Ile75(E18)Trp, Leu135(H12)Ala, and Leu135(H12)Trp (Table A6, Appendix A). Estimates of the rates of entry and escape were measured using O_2 geminate recombination as described in Scott *et al.* (*36*). The Fe atom in Mb is less reactive due to an eclipsed geometry between the edges of the coordinated imidazole ring of HisF8 and the pyrrole N atoms of the heme ring. A similar geometry occurs and reduces the reactivity of ScHbI with ligands in its T quaternary state, as mentioned in Chapter 1 (Fig. 1.7A). As a result of this geometry, very

little internal CO recombination for MbCO is observed at room temperature in ordinary buffers. Fortunately, wt Mb shows ~ 50 % geminate O_2 recombination, due to the higher reactivity of the ligand molecules, and the rate is slow enough ($k_{gem} \approx$ 10-20 μs^{-1}, $t_{1/2} \geq$ 30-60 ns) to be measured readily using a 7 ns YAG laser excitation pulse. In the case of CerHb, O_2 geminate recombination is too fast to measure accurately in the wt and most of the mutant proteins ($k_{gem} \geq 130$ μs^{-1}, $t_{1/2} \leq 5$ ns) (Chapter 5; (*74*)).

Bar graphs of the rate enhancements caused by large to small mutations, $R_{enhance}$, versus sequence position are shown in Fig. 7.4 for both SwMb and CerHb. In both proteins, significant effects are observed at the active site positions where the amino acid side chains interact directly with bound ligands (B10, CD1, E7 and E11, bars labeled with blue diamonds in Fig. 7.4). However, beyond those regions the differences between Mb and CerHb are dramatic. In the case of Mb, large rate enhancements are only observed at or near the E7 channel or in the back of the distal pocket where ligands are initially captured before binding to the Fe atom, *i.e.* positions 28(B9), 29(B10), 32(B13), 43(CD1), 46(CD4), 64(E7), and 68(E11). No significant enhancements of the rates of entry and escape are observed at positions around the Xe1 cavity, in regions equivalent to the apolar tunnel in CerHb (i.e. between the E-, G-, and H-helices), or at other locations suggested from MD simulations. In contrast, large rate enhancements for entry into and escape from CerHb are observed for replacements lining the apolar tunnel, including positions near the C-terminal ends of the E-, G-, and H-helices and at position Val7(B6) where the distal pocket connects to the apolar tunnel (bars labeled with green asterisks, Fig. 7.4B).

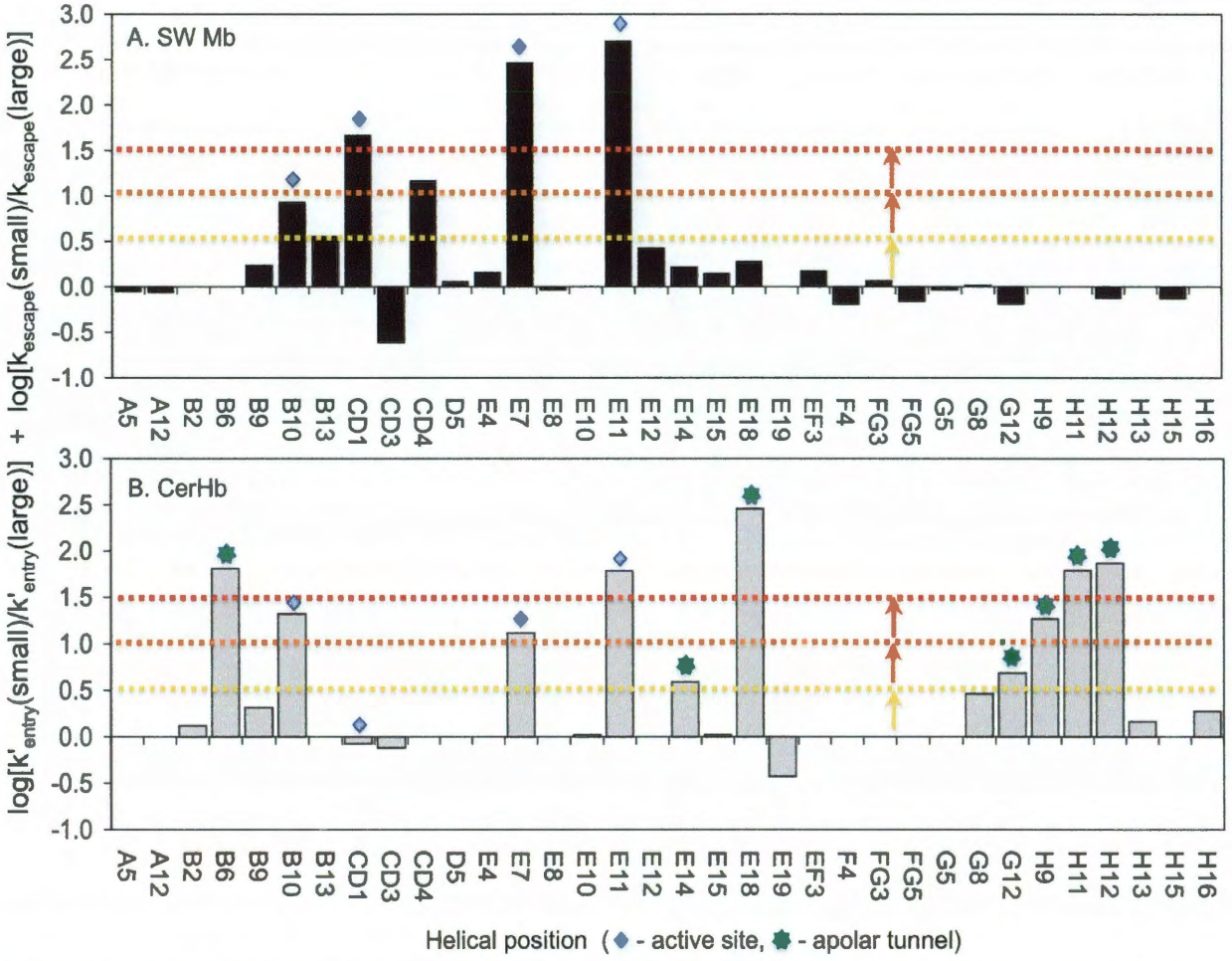


Figure 7.4. Effects of A,V to F,W mutations on entry and escape in *A*, SwMb and *B*, CerHb. Blue asterisks denote helical positions typically associated with the E7 gate pathway and green asterisks with the apolar tunnel pathway. Helical positions marked by a black diamond indicate amino acids close to or directly interacting with the coordinated ligand at the heme-binding site. Data for SwMb was taken from this study and Scott *et al.* (*36*).

7.5 Rate enhancement maps for SwMb and CerHb - The magnitudes of the $R_{enhance}$ parameters are mapped onto the three dimensional crystal structures of wt SwMb and CerHb in Fig. 7.5 (see (*36*)). The color coding of the key side chain positions is as follows: red spheres, $R_{enhance} \geq 1.5$ ($\geq$ 30-fold rate increases); orange spheres, $1.5 > R_{enhance} \geq 1.0$ (30 to10-fold increases); yellow spheres, $1.0 > R_{enhance} \geq 0.5$ (10 to 3-fold increases); white or gray sticks $0.5 > R_{enhance}$ (no significant effect, < 3-fold). The results are quite clear for both proteins. Only positions near the E7 gate or in the distal heme pocket are highlighted in Mb, and no rate enhancements are observed along alternative pathways, including the Xe pockets or other interior apolar regions. In contrast, our mutagenesis mapping approach clearly highlights the internal tunnel in CerHb, proving experimentally that ligands enter and exit this globin through this pathway. This conclusion is particularly evident in Fig. 7.5F in which the tunnel between the E- and H-helices is clearly seen and circumscribed by red, orange, and yellow side chains. The same view for Mb (Fig. 7.5E) shows that this pathway is completely blocked by the A-helix.

7.6 Experimental tests of apolar tunnel pathway in CerHb - Three further tests of the apolar pathway in CerHb were performed. First, the polar imidazole side chain of His100(H11) is extended out into solvent in the wt CerHbO_2 structures (Figs. 7.1A, 7.5B,D,F; (*74*)). As shown in Table 7.1, removing the imidazole at the H11 position by a His → Ala mutation has little affect on ligand binding, whereas both the Phe and Trp mutations cause ~ 3-fold decreases in both k'_{O2} and k_{O2} with little change in affinity. These same aromatic amino acid mutations cause large increases in the fraction of CO geminate recombination, from ~ 0.05 for wt CerHbCO to ~ 0.34 for TrpH11 (see Fig.

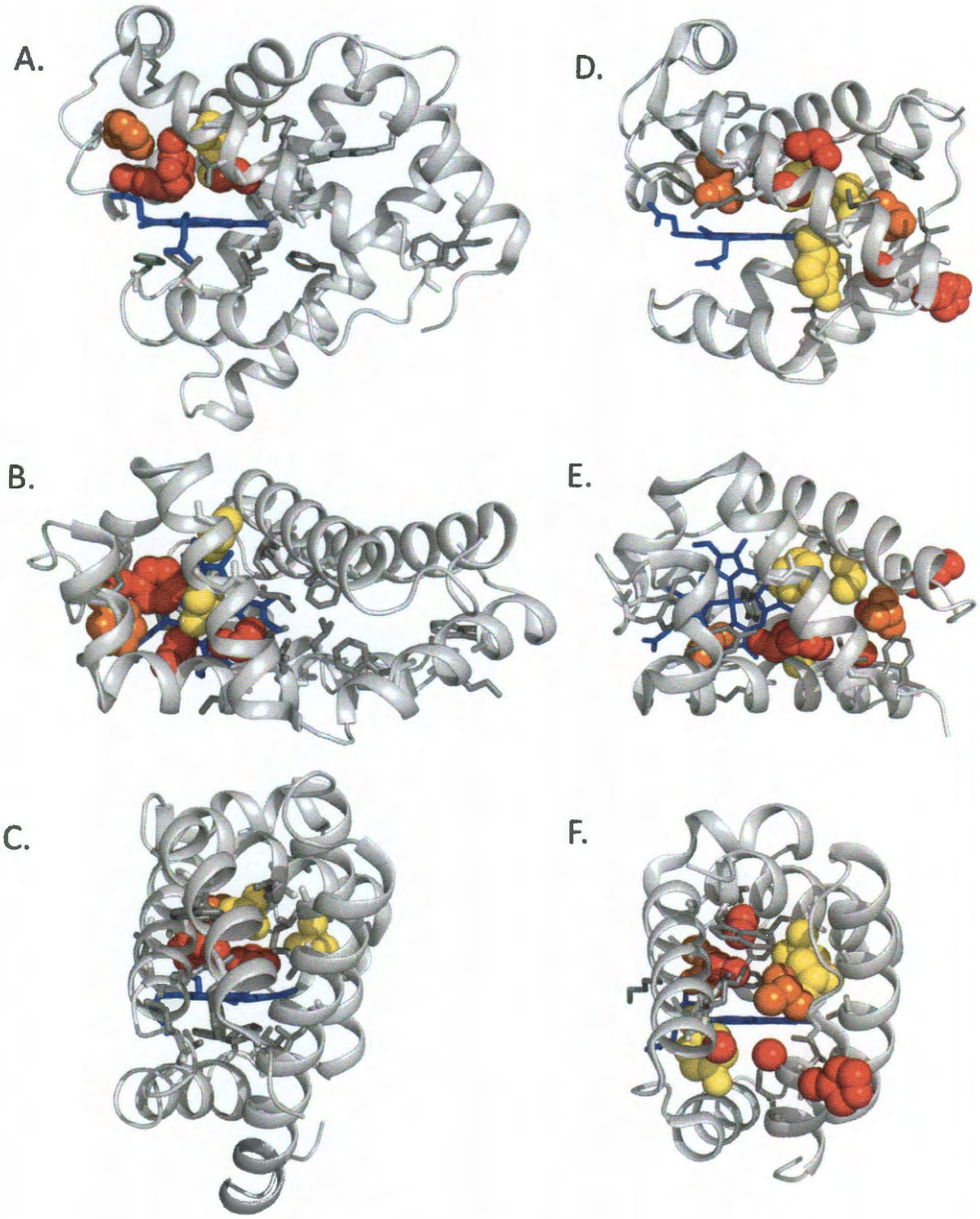


Figure 7.5. Maps of the effects of small to large mutations on the rate of Entry and Escape in SwMb (PDB: 2mgk) and CerHb (PDB: 1kr7). The protein backbones are represented as silver ribbons, the heme group is shown as blue sticks. *A*, Side view of SwMb. *B*, Side view of CerHb. *C*, Top view of SwMb. *D*, Top view of CerHb. *E*, Back view of SwMb. *F*, Tunnel/back view of CerHb. Red spheres signify a $SUM_{mutation} \geq 1.5$, orange spheres signify $1.5 \geq SUM_{mutation} \geq 1.0$, yellow spheres signify $1.0 \geq SUM_{mutation} \geq 0.5$, and gray sticks signify a $0.5 \geq SUM_{mutation}$ (numerical values can be gathered from Figure 7.4 and Scott *et al.* (*36*)). The figure was generated in MacPyMol.

7.5) and a marked decrease in $k_{gem,CO}$ due to an 8-fold decrease in k_{escape}. All of these effects suggest that the large aromatic side chains rotate back to the protein surface to partially cover the exit aperture of the ligand tunnel (Fig. 7.1, Tables 7.1 and A1, Appendix A and (*74*)).

Second, Pesce *et al.* (*42, 124*) have argued that extensive polar interactions between Tyr11(B10), Gln44(E7), Lys47(E10) and Thr48(E11) limit opening of the E7 gate, causing ligands to use the apolar tunnel for entry and exit. To test this idea, ligand binding to a triple CerHb mutant with a completely apolar active site containing PheB10/LeuE7/ValE11 mutations were examined. As shown in Table 7.1, Fig. 7.6B and Table A5 in Appendix A, rates of O_2 binding to and release from this triple mutant increase only about 50 %, and there is little change in the computed value of k'_{entry} compared to wt CerHb even though the extent of geminate recombination increases to ~ 0.40 due to a marked ~ 7-fold increase in the rate of internal bond formation compared to wt CerHb. However, when the Ala55(E18)Trp mutation is added to this triple mutant to block exit from the apolar tunnel, there are marked ~ 5-fold decreases in k'_{O2} and k_{O2}, the fraction of geminate CO rebinding increases to 0.85, and the calculated value of k'_{entry} decreases from ~ 400 to ~ 30 $\mu M^{-1}s^{-1}$ (Table 7.1 and Table A5 in Appendix A and Fig. 7.6B). Thus, even when the distal pocket is made apolar, ligands still move into and out of the protein through the apolar tunnel.

Third, to close the apolar tunnel completely, Phe and Trp mutations were constructed at: (1) the connection between the distal pocket and the interior of the tunnel, Val7(B6) (*53*); (2) in the middle of the tunnel at the Xe binding sites, Leu86(G12) (Chapter 6 and (*100*)); and (3) at the exit aperture, Ala55(E18) (Chapter 5 and (*74*)). The

Table 7.1. Rate and equilibrium parameters for O_2 CO, and NO binding to wt and selected mutants of CerHb at pH 7, 20°C. The values in parentheses in the last column are the values for k'_{NO}, the bimolecular association rate constant for NO binding, which should equal the computed rate of ligand entry, k'_{entry}.

CerHb Variant	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	$F_{gem, CO}$	k'_{entry} (k'_{NO}) $\mu M^{-1}s^{-1}$
Wt	230±17	190±25	1.2±0.2	32±5	0.05±0.03	640±400 (270)
H100A	210	230	0.93	30	0.05	570
H100F						
H100W	77	92	0.84	17	0.34	50
Y11F/Q44L/T48V	340	270	1.6	165	0.40	410
Y11F/Q44L/T48V/ A55W	70	50	1.4	27	0.85	30
Wt	230±17	190±25	1.2±0.2	32±5	0.05±0.03	640±400 (270)
V7A/L86A	240	190	1.3	35	0.14	255
V7F	110	115	0.97	22	0.16	135 (110)
L86F	120	71	1.7	24	0.17	140 (110)
A55F	61	51	1.2	17	0.44	38 (48)
V7F/L86F/A55F	42	17	2.5	17	0.79	20
V7W/L86W/A55W	23	3.0	7.6	7.2	0.93	8

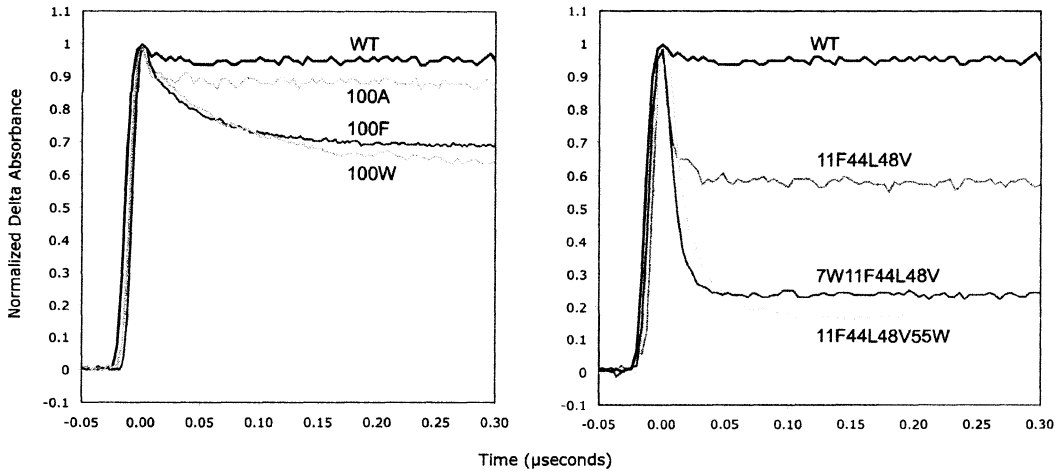


Figure 7.6. Time courses for geminate recombination in His100(H11) and apolar distal pocket mutants of CerHbCO at 20 °C, pH 7.0. *A,* CO geminate rebinding to H100A, F, and W mutants. The fraction of geminate recombination increases 6-fold upon mutating a polar His to a large nonpolar Trp. *B,* CO geminate rebinding in CerHb multiple mutants with an apolar distal pocket containing Y11F/Q44L/T48V mutations and Trp replacements at the Val7(B6) and Ala55(E18) tunnel positions.

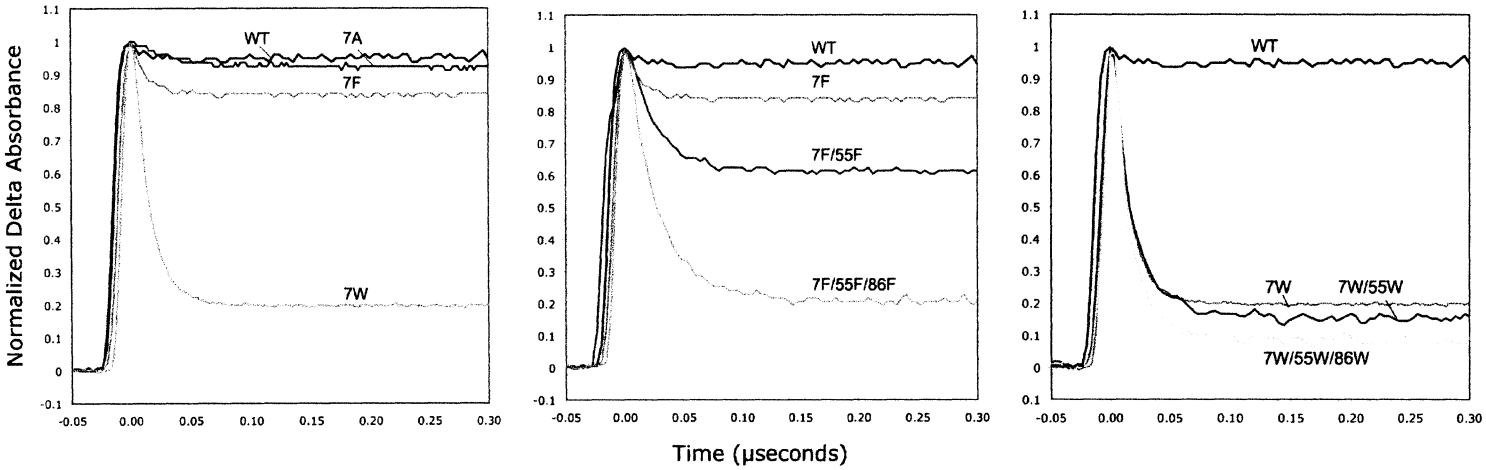


Figure 7.7. Time courses for geminate recombination in multiple tunnel mutants of CerHb at 20 °C, pH 7.0. *A*, CO geminate rebinding in position Val-7(B6) mutants of CerHb. *B*, CO geminate rebinding in single, double, and triple Phe mutants in positions Val-7(B6), Ala-55(E18), and Leu-86(G12) mutants of CerHbCO. *C*, CO geminate rebinding in single, double, and trip Trp mutants in positions Val-7(B6), Ala-55(E18), and Leu-86(G12) mutants of CerHbCO.

AlaB6/AlaG12 double mutant, which has Ala at all three positions, has kinetic parameters very similar to those of wt CerHb. In contrast, the PheB6/PheE18/PheG12 and TrpB6/TrpE18/TrpG12 triple mutants show dramatic ≥ 10-fold decreases in k'_{O2}, k_{O2}, and k'_{entry} and ≥10-fold increases in $F_{gem,CO}$, from ~ 0.05 for wt CerHb to 0.80 to 0.93 for the triple Phe and Trp mutants, respectively (Table 7.1 and Table A4 in Appendix A and Fig. 7.7). These data show that the apolar tunnel can be closed by mutagenesis, reducing the measured bimolecular rate constants for O_2 and NO binding and the calculated rate for ligand entry to less than 10 % of the wt values.

7.7 ***Single versus multiple pathways*** - The experimental mutagenesis results in Figs. 7.3-7.7 and in Table 7.1 and A1-A5, Appendix A are very clear. Ligands enter and exit CerHb through the apolar tunnel as suggested originally by Pesce *et al.* (*42*). When three Phe or Trp replacements are used to fill the entire tunnel, the bimolecular rate of ligand entry decreases dramatically. An estimate of the rate of ligand entry when the apolar channel is closed was calculated from the average of the k'_{entry}, k'_{NO}, and k'_{O2} values for the PheB6/PheE18/PheG12 and TrpB6/TrpE18/TrpG12 mutants (Tables 7.1, A1, and A5, Appendix A) and is 25 ± 12 $\mu M^{-1}s^{-1}$. An estimate of the value for ligand entry into CerHb with an open tunnel was computed from the k'_{entry}, k'_{NO}, and k'_{O2} values for wt CerHb and the AlaB6, AlaG12, and AlaB6/AlaG12 mutants and is equal to 300 ± 130 $\mu M^{-1}s^{-1}$. If the triple aromatic mutations at B6, E18, and G12 completely close the tunnel, the residual bimolecular rate of 25 $\mu M^{-1}s^{-1}$ applies to ligand binding through alternative routes. This value suggests that in wt CerHb ≥ 92 % of the ligand molecules enter through the apolar tunnel, and only 8 % by alternative routes. Interestingly, the value of k'_{entry} when the tunnel is completely blocked is only slightly smaller than the

estimated rates of ligand entry into wt and His64(E7)Gln SwMb, which are 34 and ~ 60 $\mu M^{-1}s^{-1}$, respectively. Thus, it seems likely that the limiting value of 25 $\mu M^{-1}s^{-1}$ for the triple aromatic CerHb mutants represents ligand movement through the GlnE7 gate.

At first analysis, the mapping results for SwMb seem less clear because all the positions that regulate entry and escape are located in the distal pocket or at the heme propionate solvent interface (Figs. 7.4 and 7.5). However as pointed out by Scott *et al.* (*36*) and Olson *et al.* (*32*), when the effects of Xe binding and selected mutations are examined individually, the mutagenesis results demonstrate ligand movement through the E7 gate. There are little or no changes in the fraction of geminate O_2 recombination or in the bimolecular rate constants for ligand entry, O_2, and CO binding to wt SwMb when the protein solution is pressurized with 5-10 atm of Xe (*79*). Similarly, when Trp substitutions are made in the Xe1 and Xe4 sites there are little or no changes in k'_{entry}, k_{escape} and F_{gem} (*32, 36*). The E7 gate pathway in Mb is also supported strongly and independently by time-resolved X-ray crystallography measurements. The life times of CO in the Xe1 pockets of wt, PheB10, and TrpB10 Mb are ~ 2-10 μs (*37*), 100-300 μs (*31, 35*), and 1-2 ms (*34*), respectively, and the double mutant TyrB10/GlnE7 Mb shows intermediate behavior (*125*). If ligand molecules moved directly to solvent from the Xe1 site, the size of the amino acid at the B10 position should have had no effect on ligand escape from this position. Instead, ligands appear to move back to the distal pocket, pass between the B10 side chain and the plane of the heme, and exit through the E7 channel.

When HisE7 is replaced with Trp in Mb, the calculated bimolecular rate of entry decreases from 34 μM^{-1} to 7 $\mu M^{-1}s^{-1}$, implying that up to 75 % of the ligand molecules move through the E7 gate to the active site (*36*). However, this calculation assumes that

the TrpE7 substitution completely blocks the E7 channel. Recently, Birukou *et al.* (*78*) have shown that the indole side chain in the crystal structure of TrpE7 deoxyMb does not enter the E7 channel but only partially blocks the entrance, occupying a position between the heme propionates. A similar "blocked" conformation is observed in the crystal structure of TrpE7 β human HbCO subunits (*78*). In contrast, the E7 indole side chain in the crystal structure of TrpE7 deoxy-α human Hb subunits is located in the distal pocket, completely blocking the E7 channel and sterically restricting access to the heme Fe (*78*). In this completely "closed" conformation, the α TrpE7 side chain causes ≥ 100-fold decreases in both k'_{NO} and the calculated value of k'_{entry} for isolated mutant α subunits (*73, 78*). Thus, it is possible that the percentage of ligand molecules entering through the E7 channel in Mb is significantly larger than 75 %, which was estimated from the rate of entry for TrpE7 MbO_2 where the E7 channel is not completely closed.

7.8 Discrepancies with MD simulations - The kinetic results in Fig. 7.5 are not in agreement with either past or recent molecular dynamics simulations, all of which suggest multiple (≥ 4) pathways. The latest theoretical work for both CerHb and Mb suggest 3 to 5 pathways with roughly equal probabilities, one of which is the E7 channel in Mb and one of which is the apolar tunnel in CerHb. Thus, the calculations do predict that ~ 20 to 33 % of the ligands enter and exit through the paths identified experimentally by mutagenesis. These ~ 4-fold differences compared to the experimental percentages (i.e. ~20 versus ≥ 80%) indicate that small inaccuracies in representing structural barriers in the computations could easily alter the number of plausible trajectories.

One key problem is representing the barrier to apolar ligand movement through clathrate-like water clusters that cover apolar patches at the protein surface. For example,

well-defined crystallographic water molecules surround the highly conserved Leu89(F4) side chain in Mb, which is apolar and protects the Fe-His93(F8) bond from being hydrated (*121*). Liong *et al.* (*121*) mutated this amino acid to Gly to open a hole into the Xe1 cavity, and observed that a string of 3 to 4 water molecules entered this site and were connected to well-defined surface waters, forming an extended chain of "ice-like" structures. The rates of bimolecular ligand binding and unimolecular release were relatively unaffected by this mutation, and there was little or no effect on the net extent of O_2 geminate recombination, even though the Xe1 cavity was completely open to solvent. The only major change was the loss of the slow phase of geminate rebinding, which reflects movement into and then back from the Xe1 cavity (*36*). These results show that water entered to the exposed Xe1 cavity and formed a barrier that prevented diatomic apolar ligands from entering this site. Another example of how the hydrophobic effect can influence ligand binding is the effect of the His100(H11) to Phe and Trp mutations on ligand binding to CerHb. In most published structures of wt and mutant CerHb, the imidazole side chain of HisH11 is pointing directly out into solvent (see Figs. 7.1 and 7.5). However, when this polar amino acid is replaced by Phe and Trp, the hydrophobic effect appears to push the more apolar phenyl and indole rings back on the surface of the protein, partially blocking the tunnel entrance and decreasing both k'_{entry} and k_{escape} and causing marked increases in $F_{gem,CO}$ by trapping ligands within the CerHb tunnel.

A key test of the accuracy of the theoretical algorithms would be to examine the effects of the Leu89(F4)Gly Mb and His100(H11)Trp CerHb mutations on simulated structures, starting from the wt protein, and on ligand trajectories, carefully looking at water structure in and around the Mb Xe1 cavity and the exit aperture of the CerHb

tunnel. Other valuable tests would be simulations of the effects of the ValCD4 Mb mutation on movement of the HisE7 side chain and of the motions of the indole side chain in TrpE7 Mb. The ValCD4 mutation accelerates ligand entry and release in Mb and causes outward movement of the distal histidine (*99*). The TrpE7 mutation slows ligand entry ~ 4-fold, even though only the entrance to the E7 channel is blocked (*36, 78*). The effects of all four of these mutations are clear experimentally and should be replicated in any molecular dynamics simulation. Ironically, the fastest rates of ligand uptake and release are observed when the longer, apolar tunnel trajectory is used rather than the much shorter E7 gate channel. The rates of ligand binding to CerHb are close to diffusion controlled and very similar to model heme compounds with no distal steric hindrance or to Mb and Hb mutants in which the distal histidine is replaced by Gly or Ala, creating a short, completely opened channel between solvent and the heme Fe.

7.9 Final conclusions - In our view, the experimental mutagenesis mapping results shown in Fig. 7.4 are unambiguous. Ligands enter and exit CerHb through the interior apolar tunnel at least ~ 90 % of the time, whereas ligands enter Mb from the opposite direction, ≥ 80 % of the time through the E7 gate. The same HisE7 gate pathway appears to occur in both subunits of human HbA (*73, 78*) and in the *Scapharca inaequivalvis* HbI dimer (Chapter 3 and (*81*)). The significance of these results is three-fold. (1) Our rational mutagenesis mapping strategy is robust and can identify interior apolar pathways. (2) The Mb and CerHb results provide experimental benchmarks for testing the validity of molecular dynamics trajectories, one with a short pathway between the heme propionates that is gated by a polar residue and another internal route to an external opening at the opposite end of the protein. (3) Globins have evolved different routes for

ligand entry and, the longer but more open apolar pathways allow more rapid rates of entry and exit. The HisE7 gate pathway appears to be ubiquitous in all vertebrate red cell Hbs and muscle Mbs, and probably most animal and plant Hbs with a distal histidine but otherwise apolar distal pocket (*36, 73, 78, 126*). The apolar tunnel route occurs more rarely, perhaps in response to the need for very rapid O_2 release and uptake (*43, 54, 103*).

References

1. Bolognesi, M., Bordo, D., Rizzi, M., Tarricone, C., and Ascenzi, P. (1997) Nonvertebrate hemoglobins: structural bases for reactivity, *Prog Biophys Mol Biol 68*, 29-68.
2. Wittenberg, B. A., and Wittenberg, J. B. (1989) Transport of oxygen in muscle, *Annu Rev Physiol 51*, 857-878.
3. Wittenberg, J. B. (1992) *Adv. Comp. Environ. Physiol. 13*, 60-85.
4. Kendrew, J. C. (1963) Myoglobin and the structure of proteins, *Science 139*, 1259-1266.
5. Perutz, M. F., Bolton, W., Diamond, R., Muirhead, H., and Watson, H. C. (1964) Structure of Haemoglobin. an X-Ray Examination of Reduced Horse Haemoglobin, *Nature 203*, 687-690.
6. Vinogradov, S. N., Hoogewijs, D., Bailly, X., Arredondo-Peter, R., Guertin, M., Gough, J., Dewilde, S., Moens, L., and Vanfleteren, J. R. (2005) Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life, *Proc Natl Acad Sci U S A 102*, 11385-11389.
7. Arredondo-Peter, R., Hargrove, M. S., Sarath, G., Moran, J. F., Lohrman, J., Olson, J. S., and Klucas, R. V. (1997) Rice hemoglobins. Gene cloning, analysis, and O2-binding kinetics of a recombinant protein synthesized in Escherichia coli, *Plant Physiol 115*, 1259-1266.
8. Smagghe, B. J., Hoy, J. A., Percifield, R., Kundu, S., Hargrove, M. S., Sarath, G., Hilbert, J. L., Watts, R. A., Dennis, E. S., Peacock, W. J., Dewilde, S., Moens, L., Blouin, G. C., Olson, J. S., and Appleby, C. A. (2009) Review: correlations between oxygen affinity and sequence classifications of plant hemoglobins, *Biopolymers 91*, 1083-1096.
9. Trevaskis, B., Watts, R. A., Andersson, C. R., Llewellyn, D. J., Hargrove, M. S., Olson, J. S., Dennis, E. S., and Peacock, W. J. (1997) Two hemoglobin genes in Arabidopsis thaliana: the evolutionary origins of leghemoglobins, *Proc Natl Acad Sci U S A 94*, 12230-12234.
10. Bonamore, A., and Boffi, A. (2008) Flavohemoglobin: structure and reactivity, *IUBMB Life 60*, 19-28.

11. Hankeln, T., Ebner, B., Fuchs, C., Gerlach, F., Haberkamp, M., Laufs, T. L., Roesner, A., Schmidt, M., Weich, B., Wystub, S., Saaler-Reinhardt, S., Reuss, S., Bolognesi, M., De Sanctis, D., Marden, M. C., Kiger, L., Moens, L., Dewilde, S., Nevo, E., Avivi, A., Weber, R. E., Fago, A., and Burmester, T. (2005) Neuroglobin and cytoglobin in search of their role in the vertebrate globin family, *J Inorg Biochem 99*, 110-119.

12. Pesce, A., Couture, M., Dewilde, S., Guertin, M., Yamauchi, K., Ascenzi, P., Moens, L., and Bolognesi, M. (2000) A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family, *Embo J 19*, 2424-2434.

13. Couture, M., and Guertin, M. (1996) Purification and spectroscopic characterization of a recombinant chloroplastic hemoglobin from the green unicellular alga Chlamydomonas eugametos, *Eur J Biochem 242*, 779-787.

14. Watts, R. A., Hunt, P. W., Hvitved, A. N., Hargrove, M. S., Peacock, W. J., and Dennis, E. S. (2001) A hemoglobin from plants homologous to truncated hemoglobins of microorganisms, *Proc Natl Acad Sci U S A 98*, 10119-10124.

15. Wittenberg, J. B., Bolognesi, M., Wittenberg, B. A., and Guertin, M. (2002) Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants, *J Biol Chem 277*, 871-874.

16. Milani, M., Pesce, A., Nardini, M., Ouellet, H., Ouellet, Y., Dewilde, S., Bocedi, A., Ascenzi, P., Guertin, M., Moens, L., Friedman, J. M., Wittenberg, J. B., and Bolognesi, M. (2005) Structural bases for heme binding and diatomic ligand recognition in truncated hemoglobins, *J Inorg Biochem 99*, 97-109.

17. Brunori, M. (2001) Nitric oxide moves myoglobin centre stage, *Trends Biochem Sci 26*, 209-210.

18. Brunori, M. (2001) Nitric oxide, cytochrome-c oxidase and myoglobin, *Trends Biochem Sci 26*, 21-23.

19. Dou, Y., Maillett, D. H., Eich, R. F., and Olson, J. S. (2002) Myoglobin as a model system for designing heme protein based blood substitutes, *Biophys Chem 98*, 127-148.

20. Eich, R. F., Li, T., Lemon, D. D., Doherty, D. H., Curry, S. R., Aitken, J. F., Mathews, A. J., Johnson, K. A., Smith, R. D., Phillips, G. N., Jr., and Olson, J. S. (1996) Mechanism of NO-induced oxidation of myoglobin and hemoglobin, *Biochemistry 35*, 6976-6983.

21. Herold, S., Exner, M., and Nauser, T. (2001) Kinetic and mechanistic studies of the NO*-mediated oxidation of oxymyoglobin and oxyhemoglobin, *Biochemistry 40*, 3385-3395.

22. Gardner, P. R., Costantino, G., Szabo, C., and Salzman, A. L. (1997) Nitric oxide sensitivity of the aconitases, *J Biol Chem 272*, 25071-25076.

23. Gardner, P. R. (2005) Nitric oxide dioxygenase function and mechanism of flavohemoglobin, hemoglobin, myoglobin and their associated reductases, *J Inorg Biochem 99*, 247-266.

24. Gardner, P. R., Gardner, A. M., Brashear, W. T., Suzuki, T., Hvitved, A. N., Setchell, K. D., and Olson, J. S. (2006) Hemoglobins dioxygenate nitric oxide with high fidelity, *J Inorg Biochem 100*, 542-550.

25. Herold, S. (1999) Kinetic and spectroscopic characterization of an intermediate peroxynitrite complex in the nitrogen monoxide induced oxidation of oxyhemoglobin, *FEBS Lett 443*, 81-84.

26. Hou, S., Larsen, R. W., Boudko, D., Riley, C. W., Karatan, E., Zimmer, M., Ordal, G. W., and Alam, M. (2000) Myoglobin-like aerotaxis transducers in Archaea and Bacteria, *Nature 403*, 540-544.

27. LaCount, M. W., Zhang, E., Chen, Y. P., Han, K., Whitton, M. M., Lincoln, D. E., Woodin, S. A., and Lebioda, L. (2000) The crystal structure and amino acid sequence of dehaloperoxidase from Amphitrite ornata indicate common ancestry with globins, *J Biol Chem 275*, 18712-18716.

28. Lebioda, L., LaCount, M. W., Zhang, E., Chen, Y. P., Han, K., Whitton, M. M., Lincoln, D. E., and Woodin, S. A. (1999) An enzymatic globin from a marine worm, *Nature 401*, 445.

29. Weber, R. E., and Vinogradov, S. N. (2001) Nonvertebrate hemoglobins: functions and molecular adaptations, *Physiol Rev 81*, 569-628.

30. Perutz, M. F., and Mathews, F. S. (1966) An x-ray study of azide methaemoglobin, *J Mol Biol 21*, 199-202.

31. Aranda, R. t., Levin, E. J., Schotte, F., Anfinrud, P. A., and Phillips, G. N., Jr. (2006) Time-dependent atomic coordinates for the dissociation of carbon monoxide from myoglobin, *Acta Crystallogr D Biol Crystallogr 62*, 776-783.

32. Olson, J. S., Soman, J., and Phillips, G. N., Jr. (2007) Ligand pathways in myoglobin: A review of trp cavity mutations, *IUBMB Life 59*, 552-562.

33. Olson, J. S. G., Abhik (2008) Mammalian Myoglobin as a Model for Understanding Ligand Affinities and Discrimination in Heme Proteins., In *The Smallest Biomolecules: Perspectives on Heme-Diatomic Interactions* (Ghosh, A., Ed.), pp 3-17, Elsevier, London.

34. Schmidt, M., Nienhaus, K., Pahl, R., Krasselt, A., Anderson, S., Parak, F., Nienhaus, G. U., and Srajer, V. (2005) Ligand migration pathway and protein dynamics in myoglobin: a time-resolved crystallographic study on L29W MbCO, *Proc Natl Acad Sci U S A 102*, 11704-11709.

35. Schotte, F., Soman, J., Olson, J. S., Wulff, M., and Anfinrud, P. A. (2004) Picosecond time-resolved X-ray crystallography: probing protein function in real time, *J Struct Biol 147*, 235-246.

36. Scott, E. E., Gibson, Q. H., and Olson, J. S. (2001) Mapping the pathways for O2 entry into and exit from myoglobin, *J Biol Chem 276*, 5177-5188.

37. Srajer, V., Ren, Z., Teng, T. Y., Schmidt, M., Ursby, T., Bourgeois, D., Pradervand, C., Schildkamp, W., Wulff, M., and Moffat, K. (2001) Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction, *Biochemistry 40*, 13802-13815.

38. Olson, D. J. S. (2008) From O2 diffusion into red blood cells to ligand pathways in globins, In *Dioxygen binding and sensing proteins* (Bolognesi, M., di Prisco, G., and Verde, C., Eds.), Springer-Verlag Italia Srl, Milano, Italy.

39. Elber, R. (2010) Ligand diffusion in globins: simulations versus experiment, *Curr Opin Struct Biol 20*, 162-167.

40. Pardanani, A., Gambacurta, A., Ascoli, F., and Royer, W. E., Jr. (1998) Mutational destabilization of the critical interface water cluster in Scapharca dimeric hemoglobin: structural basis for altered allosteric activity, *J Mol Biol 284*, 729-739.

41. Royer, W. E., Jr. (1994) High-resolution crystallographic analysis of a co-operative dimeric hemoglobin, *J Mol Biol 235*, 657-681.

42. Pesce, A., Nardini, M., Dewilde, S., Geuens, E., Yamauchi, K., Ascenzi, P., Riggs, A. F., Moens, L., and Bolognesi, M. (2002) The 109 residue nerve tissue minihemoglobin from Cerebratulus lacteus highlights striking structural plasticity of the alpha-helical globin fold, *Structure 10*, 725-735.

43. Vandergon, T. L., Riggs, C. K., Gorr, T. A., Colacino, J. M., and Riggs, A. F. (1998) The mini-hemoglobins in neural and body wall tissue of the nemertean worm, Cerebratulus lacteus, *J Biol Chem 273*, 16998-17011.

44. Royer, W. E., Jr., Knapp, J. E., Strand, K., and Heaslet, H. A. (2001) Cooperative hemoglobins: conserved fold, diverse quaternary assemblies and allosteric mechanisms, *Trends Biochem Sci 26*, 297-304.

45. Royer, W. E., Jr., Zhu, H., Gorr, T. A., Flores, J. F., and Knapp, J. E. (2005) Allosteric hemoglobin assembly: diversity and similarity, *J Biol Chem 280*, 27477-27480.

46. Summerford, C. M., Pardanani, A., Betts, A. H., Poteete, A. R., Colotti, G., and Royer, W. E., Jr. (1995) Bacterial expression of Scapharca dimeric hemoglobin: a simple model system for investigating protein cooperatively, *Protein Eng 8*, 593-599.

47. Nienhaus, K., Knapp, J. E., Palladino, P., Royer, W. E., Jr., and Nienhaus, G. U. (2007) Ligand migration and binding in the dimeric hemoglobin of Scapharca inaequivalvis, *Biochemistry 46*, 14018-14031.

48. Royer, W. E., Jr., Heard, K. S., Harrington, D. J., and Chiancone, E. (1995) The 2.0 A crystal structure of Scapharca tetrameric hemoglobin: cooperative dimers within an allosteric tetramer, *J Mol Biol 253*, 168-186.

49. Ceci, P., Giangiacomo, L., Boffi, A., and Chiancone, E. (2002) The mutation K30D disrupts the only salt bridge at the subunit interface of the homodimeric

hemoglobin from Scapharca inaequivalvis and changes the mechanism of cooperativity, *J Biol Chem 277*, 6929-6933.

50. Guarrera, L., Colotti, G., Boffi, A., Chiancone, E., Das, T. K., Rousseau, D. L., and Gibson, Q. H. (1998) The apolar distal histidine mutant (His69-->Val) of the homodimeric Scapharca hemoglobin is in an R-like conformation, *Biochemistry 37*, 5608-5615.

51. Knapp, J. E., Bonham, M. A., Gibson, Q. H., Nichols, J. C., and Royer, W. E., Jr. (2005) Residue F4 plays a key role in modulating oxygen affinity and cooperativity in Scapharca dimeric hemoglobin, *Biochemistry 44*, 14419-14430.

52. Bidon-Chanal, A., Marti, M. A., Crespo, A., Milani, M., Orozco, M., Bolognesi, M., Luque, F. J., and Estrin, D. A. (2006) Ligand-induced dynamical regulation of NO conversion in Mycobacterium tuberculosis truncated hemoglobin-N, *Proteins 64*, 457-464.

53. Deng, P., Nienhaus, K., Palladino, P., Olson, J. S., Blouin, G., Moens, L., Dewilde, S., Geuens, E., and Nienhaus, G. U. (2007) Transient ligand docking sites in Cerebratulus lacteus mini-hemoglobin, *Gene 398*, 208-223.

54. Milani, M., Pesce, A., Ouellet, Y., Ascenzi, P., Guertin, M., and Bolognesi, M. (2001) Mycobacterium tuberculosis hemoglobin N displays a protein tunnel suited for O2 diffusion to the heme, *Embo J 20*, 3902-3909.

55. Bidon-Chanal, A., Marti, M. A., Estrin, D. A., and Luque, F. J. (2007) Dynamical regulation of ligand migration by a gate-opening molecular switch in truncated hemoglobin-N from Mycobacterium tuberculosis, *J Am Chem Soc 129*, 6782-6788.

56. Crespo, A., Marti, M. A., Kalko, S. G., Morreale, A., Orozco, M., Gelpi, J. L., Luque, F. J., and Estrin, D. A. (2005) Theoretical study of the truncated hemoglobin HbN: exploring the molecular basis of the NO detoxification mechanism, *J Am Chem Soc 127*, 4433-4444.

57. Milani, M., Pesce, A., Ouellet, H., Guertin, M., and Bolognesi, M. (2003) Truncated hemoglobins and nitric oxide action, *IUBMB Life 55*, 623-627.

58. Ouellet, Y. H., Daigle, R., Lague, P., Dantsker, D., Milani, M., Bolognesi, M., Friedman, J. M., and Guertin, M. (2008) Ligand binding to truncated hemoglobin

N from Mycobacterium tuberculosis is strongly modulated by the interplay between the distal heme pocket residues and internal water, *J Biol Chem 283*, 27270-27278.

59. Guallar, V., Lu, C., Borrelli, K., Egawa, T., and Yeh, S. R. (2009) Ligand migration in the truncated hemoglobin-II from Mycobacterium tuberculosis: the role of G8 tryptophan, *J Biol Chem 284*, 3106-3116.

60. Pesce, A., Nardini, M., Dewilde, S., Ascenzi, P., Riggs, A. F., Yamauchi, K., Geuens, E., Moens, L., and Bolognesi, M. (2001) Crystallization and preliminary X-ray analysis of neural haemoglobin from the nemertean worm Cerebratulus lacteus, *Acta Crystallogr D Biol Crystallogr 57*, 1897-1899.

61. Draghi, F., Miele, A. E., Travaglini-Allocatelli, C., Vallone, B., Brunori, M., Gibson, Q. H., and Olson, J. S. (2002) Controlling ligand binding in myoglobin by mutagenesis, *J Biol Chem 277*, 7509-7519.

62. Pesce, A., Nardini, M., Ascenzi, P., Geuens, E., Dewilde, S., Moens, L., Bolognesi, M., Riggs, A. F., Hale, A., Deng, P., Nienhaus, G. U., Olson, J. S., and Nienhaus, K. (2004) Thr-E11 regulates O2 affinity in Cerebratulus lacteus mini-hemoglobin, *J Biol Chem 279*, 33662-33672.

63. Brent, R., and Ptashne, M. (1981) Mechanism of action of the lexA gene product, *Proc Natl Acad Sci U S A 78*, 4204-4208.

64. Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F., and Antonini, E. (1981) Dimeric and tetrameric hemoglobins from the mollusc Scapharca inaequivalvis. Structural and functional properties, *J Mol Biol 152*, 577-592.

65. Pardanani, A., Gibson, Q. H., Colotti, G., and Royer, W. E., Jr. (1997) Mutation of residue Phe97 to Leu disrupts the central allosteric pathway in Scapharca dimeric hemoglobin, *J Biol Chem 272*, 13171-13179.

66. Das, T. K., Weber, R. E., Dewilde, S., Wittenberg, J. B., Wittenberg, B. A., Yamauchi, K., Van Hauwaert, M. L., Moens, L., and Rousseau, D. L. (2000) Ligand binding in the ferric and ferrous states of Paramecium hemoglobin, *Biochemistry 39*, 14330-14340.

67. Dewilde, S., Blaxter, M., Van Hauwaert, M. L., Van Houte, K., Pesce, A., Griffon, N., Kiger, L., Marden, M. C., Vermeire, S., Vanfleteren, J., Esmans, E.,

and Moens, L. (1998) Structural, functional, and genetic characterization of Gastrophilus hemoglobin, *J Biol Chem 273*, 32467-32474.

68. Ikehara, M., Ohtsuka, E., Tokunaga, T., Taniyama, Y., Iwai, S., Kitano, K., Miyamoto, S., Ohgi, T., Sakuragawa, Y., Fujiyama, K., and et al. (1984) Synthesis of a gene for human growth hormone and its expression in Escherichia coli, *Proc Natl Acad Sci U S A 81*, 5956-5960.

69. Springer, B. A., and Sligar, S. G. (1987) High-level expression of sperm whale myoglobin in Escherichia coli, *Proc Natl Acad Sci U S A 84*, 8961-8965.

70. Carver, T. E., Brantley, R. E., Jr., Singleton, E. W., Arduini, R. M., Quillin, M. L., Phillips, G. N., Jr., and Olson, J. S. (1992) A novel site-directed mutant of myoglobin with an unusually high O2 affinity and low autooxidation rate, *J Biol Chem 267*, 14443-14450.

71. Olson, J. S., Foley, E. W., Maillett, D. H., and Paster, E. V. (2003) Measurement of rate constants for reactions of O2, CO, and NO with hemoglobin, *Methods Mol Med 82*, 65-91.

72. Rohlfs, R. J., Mathews, A. J., Carver, T. E., Olson, J. S., Springer, B. A., Egeberg, K. D., and Sligar, S. G. (1990) The effects of amino acid substitution at position E7 (residue 64) on the kinetics of ligand binding to sperm whale myoglobin, *J Biol Chem 265*, 3168-3176.

73. Birukou, I., Schweers, R. L., and Olson, J. S. (2010) Distal histidine stabilizes bound O2 and acts as a gate for ligand entry in both subunits of adult human hemoglobin, *J Biol Chem 285*, 8840-8854.

74. Salter, M. D., Nienhaus, K., Nienhaus, G. U., Dewilde, S., Moens, L., Pesce, A., Nardini, M., Bolognesi, M., and Olson, J. S. (2008) The apolar channel in Cerebratulus lacteus hemoglobin is the route for O2 entry and exit, *J Biol Chem 283*, 35689-35702.

75. Schweers, R. L. (2003) Electrostatic regulation of O2 and CO binding in the alpha and beta subunits of recombinant human hemoglobin, In *Biochemistry and Cell Biology*, Rice University, Houston.

76. Olson, D. J. S. (1981) In *Methods Enzymol*, pp 631-651.

77. Blouin, G. C. (2008) Alkyl isocyanids as transition state analogs for ligand entry and exit in globins, In *Biochemistry and Cell Biology*, p 207, Rice University, Houston.

78. Birukou, I., Soman, J., and Olson, J. S. (2010) Blocking the gate to ligand entry in human hemoglobin, *J Biol Chem.*

79. Scott, E. E., and Gibson, Q. H. (1997) Ligand migration in sperm whale myoglobin, *Biochemistry 36*, 11909-11917.

80. Chiancone, E., Elber, R., Royer, W. E., Jr., Regan, R., and Gibson, Q. H. (1993) Ligand binding and conformation change in the dimeric hemoglobin of the clam Scapharca inaequivalvis, *J Biol Chem 268*, 5711-5718.

81. Knapp, J. E., Pahl, R., Cohen, J., Nichols, J. C., Schulten, K., Gibson, Q. H., Srajer, V., and Royer, W. E., Jr. (2009) Ligand migration and cavities within Scapharca Dimeric HbI: studies by time-resolved crystallo-graphy, Xe binding, and computational analysis, *Structure 17*, 1494-1504.

82. Li, T., Quillin, M. L., Phillips, G. N., Jr., and Olson, J. S. (1994) Structural determinants of the stretching frequency of CO bound to myoglobin, *Biochemistry 33*, 1433-1446.

83. Birukou, I., Soman, J., and Olson, J. S. (2011) Blocking the gate to ligand entry in human hemoglobin, *J Biol Chem 286*, 10515.

84. Olson, J. S., and Phillips, G. N., Jr. (1996) Kinetic pathways and barriers for ligand binding to myoglobin, *J Biol Chem 271*, 17593-17596.

85. Phillips, G. N., Jr., Teodoro, M., Li, T., Smith, B., Gilson, M.M. and Olson, J.S. (1999) Bound CO is a molecular probe of electrostatic potential in the distal pocket of myoglobin, *J. Phys. Chem. B. 103*, 8817-8829.

86. Deinum, G., Stone, J. R., Babcock, G. T., and Marletta, M. A. (1996) Binding of nitric oxide and carbon monoxide to soluble guanylate cyclase as observed with Resonance raman spectroscopy, *Biochemistry 35*, 1540-1547.

87. Mitchell, D. M., Shapleigh, J. P., Archer, A. M., Alben, J. O., and Gennis, R. B. (1996) A pH-dependent polarity change at the binuclear center of reduced cytochrome c oxidase detected by FTIR difference spectroscopy of the CO adduct, *Biochemistry 35*, 9446-9450.

88. Ray, G. B., Li, X.Y., Ibers, J.A., Sessler, J.L., Spiro, T.G. (1994) How far can proteins bend the FeCO unit? Distal polar and steric effects in heme proteins and models, *J. Amer. Chem. Soc. 116*, 162-176.

89. Olson, D. J. S., and Phillips, G. N., Jr. (1997) Myoglobin discriminates between O2, NO, and CO by electrostatic interactions with the bound ligand, *J Biol Inorg Chem 2*, 544-552.

90. Hale, A. D. (2004) Electrostatic regulation of oxygen binding to the neuronal hemoglobin of Cerebratulus lacteus, In *Biochemistry and Cell Biology*, p 42, Rice University, Houston.

91. Hirota, S., Li, T. S., Phillips, G. N. J., Olson, J. S., and Kitagawa, T. (1996) Perturbation of the Fe-O-2 bond by nearby residues in heme pocket - observation of nu(Fe-O2) Raman bands for oxymyoglobin mutants., *J. Amer. Chem. Soc. 118*, 7845-7846.

92. Kundu, S., Snyder, B., Das, K., Chowdhury, P., Park, J., Petrich, J. W., and Hargrove, M. S. (2002) The leghemoglobin proximal heme pocket directs oxygen dissociation and stabilizes bound heme, *Proteins 46*, 268-277.

93. Springer, B. A., Sligar, S.G., Olson, J.S. and Phillips, G.N., Jr. (1994) Mechanisms of ligand recognition in myoglobin, *Chem Rev 94*, 699-714.

94. Huang, S., Huang, J., Kloek, A. P., Goldberg, D. E., and Friedman, J. M. (1996) Hydrogen bonding of tyrosine B10 to heme-bound oxygen in Ascaris hemoglobin. Direct evidence from UV resonance Raman spectroscopy, *J Biol Chem 271*, 958-962.

95. Kloek, A. P., Yang, J., Mathews, F. S., Frieden, C., and Goldberg, D. E. (1994) The tyrosine B10 hydroxyl is crucial for oxygen avidity of Ascaris hemoglobin, *J Biol Chem 269*, 2377-2379.

96. Carver, T. E., Rohlfs, R. J., Olson, J. S., Gibson, Q. H., Blackmore, R. S., Springer, B. A., and Sligar, S. G. (1990) Analysis of the kinetic barriers for ligand binding to sperm whale myoglobin using site-directed mutagenesis and laser photolysis techniques, *J Biol Chem 265*, 20007-20020.

97. Laskowski, R. A. (1995) SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions, *J Mol Graph 13*, 323-330, 307-328.

98. Yang, F., and Phillips, G. N., Jr. (1996) Crystal structures of CO-, deoxy- and met-myoglobins at various pH values, *J Mol Biol 256*, 762-774.

99. Lai, H. H., Li, T., Lyons, D. S., Phillips, G. N., Jr., Olson, J. S., and Gibson, Q. H. (1995) Phe-46(CD4) orients the distal histidine for hydrogen bonding to bound ligands in sperm whale myoglobin, *Proteins 22*, 322-339.

100. Pesce, A., Nardini, M., Dewilde, S., Capece, L., Marti, M. A., Congia, S., Salter, M. D., Blouin, G. C., Estrin, D. A., Ascenzi, P., Moens, L., Bolognesi, M., and Olson, J. S. (2011) Ligand migration in the apolar tunnel of cerebratulus lacteus mini-hemoglobin, *J Biol Chem 286*, 5347-5358.

101. Perutz, M. F. (1979) Regulation of oxygen affinity of hemoglobin: influence of structure of the globin on the heme iron, *Annu Rev Biochem 48*, 327-386.

102. de Sanctis, D., Dewilde, S., Pesce, A., Moens, L., Ascenzi, P., Hankeln, T., Burmester, T., and Bolognesi, M. (2004) Mapping protein matrix cavities in human cytoglobin through Xe atom binding, *Biochem Biophys Res Commun 316*, 1217-1221.

103. Milani, M., Pesce, A., Ouellet, Y., Dewilde, S., Friedman, J., Ascenzi, P., Guertin, M., and Bolognesi, M. (2004) Heme-ligand tunneling in group I truncated hemoglobins, *J Biol Chem 279*, 21520-21525.

104. Raushel, F. M., Thoden, J. B., and Holden, H. M. (2003) Enzymes with molecular tunnels, *Acc Chem Res 36*, 539-548.

105. Moschetti, T., Mueller, U., Schulze, J., Brunori, M., and Vallone, B. (2009) The structure of neuroglobin at high Xe and Kr pressure reveals partial conservation of globin internal cavities, *Biophys J 97*, 1700-1708.

106. Savino, C., Miele, A. E., Draghi, F., Johnson, K. A., Sciara, G., Brunori, M., and Vallone, B. (2009) Pattern of cavities in globins: the case of human hemoglobin, *Biopolymers 91*, 1097-1107.

107. Schoenborn, B. P. (1965) Binding of xenon to horse haemoglobin, *Nature 208*, 760-762.

108. Schoenborn, B. P., Watson, H. C., and Kendrew, J. C. (1965) Binding of xenon to sperm whale myoglobin, *Nature 207*, 28-30.

109. Tilton, R. F., Jr., Kuntz, I. D., Jr., and Petsko, G. A. (1984) Cavities in proteins: structure of a metmyoglobin-xenon complex solved to 1.9 A, *Biochemistry 23*, 2849-2857.

110. Elber, R., and Karplus, M. (1990) Enhanced sampling in molecular dynamics: use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin, *J. Am. Chem. Soc. 112*, 9161-9175.

111. Bossa, C., Amadei, A., Daidone, I., Anselmi, M., Vallone, B., Brunori, M., and Di Nola, A. (2005) Molecular dynamics simulation of sperm whale myoglobin: effects of mutations and trapped CO on the structure and dynamics of cavities, *Biophys J 89*, 465-474.

112. Brunori, M., Bourgeois, D., and Vallone, B. (2004) The structural dynamics of myoglobin, *J Struct Biol 147*, 223-234.

113. Elber, R., and Gibson, Q. H. (2008) Toward quantitative simulations of carbon monoxide escape pathways in myoglobin, *J Phys Chem B 112*, 6147-6154.

114. Golden, S. D., and Olsen, K. W. (2008) Use of the conjugate peak refinement algorithm for identification of ligand-binding pathways in globins, *Methods Enzymol 437*, 417-437.

115. Mouawad, L., Marechal, J. D., and Perahia, D. (2005) Internal cavities and ligand passageways in human hemoglobin characterized by molecular dynamics simulations, *Biochim Biophys Acta 1724*, 385-393.

116. Nutt, D. R., and Meuwly, M. (2004) CO migration in native and mutant myoglobin: atomistic simulations for the understanding of protein function, *Proc Natl Acad Sci U S A 101*, 5998-6002.

117. Orlowski, S., and Nowak, W. (2007) Locally enhanced sampling molecular dynamics study of the dioxygen transport in human cytoglobin, *J Mol Model 13*, 715-723.

118. Orlowski, S., and Nowak, W. (2008) Topology and thermodynamics of gaseous ligands diffusion paths in human neuroglobin, *Biosystems 94*, 263-266.

119. Ruscio, J. Z., Kumar, D., Shukla, M., Prisant, M. G., Murali, T. M., and Onufriev, A. V. (2008) Atomic level computational identification of ligand migration pathways between solvent and binding site in myoglobin, *Proc Natl Acad Sci U S A 105*, 9204-9209.

120. Huang, X., and Boxer, S. G. (1994) Discovery of new ligand binding pathways in myoglobin by random mutagenesis, *Nat Struct Biol 1*, 226-229.

121. Liong, E. C., Dou, Y., Scott, E. E., Olson, J. S., and Phillips, G. N., Jr. (2001) Waterproofing the heme pocket. Role of proximal amino acid side chains in preventing hemin loss from myoglobin, *J Biol Chem 276*, 9093-9100.

122. Quillin, M. L., Arduini, R. M., Olson, J. S., and Phillips, G. N., Jr. (1993) High-resolution crystal structures of distal histidine mutants of sperm whale myoglobin, *J Mol Biol 234*, 140-155.

123. Quillin, M. L., Li, T., Olson, J. S., Phillips, G. N., Jr., Dou, Y., Ikeda-Saito, M., Regan, R., Carlson, M., Gibson, Q. H., Li, H., and et al. (1995) Structural and functional effects of apolar mutations of the distal valine in myoglobin, *J Mol Biol 245*, 416-436.

124. Pesce, A., Nardini, M., Dewilde, S., Capece, L., Marti, M. A., Congia, S., Salter, M. D., Blouin, G. C., Estrin, D. A., Ascenzi, P., Moens, L., Bolognesi, M., and Olson, J. S. (2010) Ligand migration in the apolar tunnel of cerebratulus lacteus mini-hemoglobin, *J Biol Chem*.

125. Bourgeois, D., Vallone, B., Arcovito, A., Sciara, G., Schotte, F., Anfinrud, P. A., and Brunori, M. (2006) Extended subnanosecond structural dynamics of myoglobin revealed by Laue crystallography, *Proc Natl Acad Sci U S A 103*, 4924-4929.

126. Kundu, S., and Hargrove, M. S. (2003) Distal heme pocket regulation of ligand binding and stability in soybean leghemoglobin, *Proteins 50*, 239-248.

Appendix A

Table A1. Bimolecular, geminate, and calculated rate parameters for CerHb mutants located throughout the apolar tunnel measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.0053 (15)	590±10 0	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Trp3(B2)											
Ala3	220	140	1.6	33	0.060	550	140	0.068	9.5	130	490
Leu3	260	190	1.4	22	0.056	390	54	0.040	2.2	52	550
Phe3	260	180	1.4	25	0.047	530	92	0.069	6.3	86	360
Tyr3	230	160	1.4	26	0.060	430	52	0.040	2.1	50	650
Val7(B6)											
Ala7	250	270	0.93	36	0.033	1,100	33	0.080	2.6	30	450
Phe7	110	110	1.0	22	0.027	820	64	0.16	10	54	140
Trp7	33	14	2.4	13	0.0042	3,100	70	0.82	57	13	16
Tyr51(E14)											
Ala51	260	150	1.7	38	0.015 (26%) 0.17 (74%)	2,500 230	200 15	0.16	32 2.4	90[b]	240
Leu51	370	500	0.74	39	0.039	1,000	30	0.16	4.8	25	240
Phe51	270	280	0.96	37	0.041	900	96	0.087	8.4	88	430
Trp51	130	89	1.5	38	0.020	1,900	50	0.26	13	37	150
Ile52(E15)											
Ala52	130	75	1.7	31	0.036	860	12	0.28	3.4	9.0	110
Phe52	67	74	0.91	19	0.022	860	87	0.30	26	61	63
Trp52	47	25	1.9	11	0.016	690	59	0.41	24	35	27
Ala55(E18)											
Val55	180	93	1.9	41	0.045	910	24	0.16	3.8	20	260
Leu55	72	40	1.8	17	0.016	1,100	7.6	0.46	3.5	4.0	37
Phe55	61	51	1.2	17	0.029	590	9.3	0.44	4.1	5.0	39
Trp55	63	66	0.95	15	0.029	520	9.4	0.44	4.1	5.0	34

Table A1 (cont'd). Bimolecular, geminate, and calculated rate parameters for CerHb mutants located throughout the apolar tunnel measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.0053 (15)	590±10 0	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Ile56(E19)											
Ala56	180	94	1.9	17	0.056	300	40	0.15	6.0	34	110
Phe56	200	130	1.5	35	0.055	640	58	0.11	6.4	52	320
Trp56	170	110	1.5	33	0.073	450	46	0.13	6.0	40	250
Leu86(G12)											
Ala86	230	150	1.5	25	0.053	470	23	0.10	2.3	21	250
Phe86	120	71	1.7	24	0.024	1,000	92	0.16	15	77	150
Trp86	97	60	1.6	16	0.014	1,100	78	0.54	42	36	30
Leu98(H9)											
Ala98	150	160	0.93	18	0.059	300	200	0.15	30	170	120
Phe98	120	74	1.6	23	0.038	600	61	0.11	6.7	54	210
Trp98	61	62	0.98	15	0.032	470	31	0.30	9.3	22	50
His100(H11)											
Ala100	210	230	0.91	30	0.056	540	80	0.10	8.0	72	300
Phe100	99	110	0.90	17	0.032 0.00030	530 57,000	20	0.29	5.8	14	59
Trp100	77	92	0.84	17	0.036	470	11	0.34	3.7	7.0	50
Ala101(H12)											
Phe101	77	68	1.1	16	0.027	590	8.1	0.38	3.1	5.0	42
Trp101	100	65	1.5	23	0.035	660	11	0.28	3.1	8.0	82

Table A1 (cont'd). Bimolecular, geminate, and calculated rate parameters for CerHb mutants located throughout the apolar tunnel measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.0053 (15)	590±10 0	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Ile102(H13)											
Ala102	270	190	1.4	35	0.036	970	19	0.12	2.3	17	290
Phe102	190	120	1.6	23	0.040	580	25	0.060	1.5	24	380
Trp102	120	60	2.0	24	0.054 (60%)	440	32	0.19	6.1	26	130
					0.0069 (40%)	3,500					
Asp104(H15)											
Val104	230	170	1.4	36	0.063	570	35	0.072	2.5	32	500
Leu104	270	200	1.4	52	0.050	1,000	44	0.15	6.6	37	350
Ile105(H16)											
Ala105	290	98	3.0	80	0.026	3,100	39	0.23	9.0	30	350
Phe105	230	260	0.88	22	0.025	880	100	0.15	15	85	150
					0.00030	73,000					
Trp105	160	64	2.5	32	0.014	2,300	73	0.30	22	51	110
					0.0020	16,000					

The standard deviations for the wt values for k'_{O2}, k_{O2}, and k'_{CO} were obtained from the analysis of 12 completely independent sets of experiments starting with at least 7 different expressions and purifications of CerHb. The standard deviation for k_{CO} came from the analysis of 16 different determinations. The values of K_{O2} and K_{CO} were calculated from k'_{O2}/k_{O2} and k'_{CO}/k_{CO}, and their standard deviations were calculated from the standard propagation of error formula.

[a] The fraction of geminate CO rebinding was measured using a 7 ns YAG laser light pulse.

[b] k_{escape} calculated from 50 % of each k_{gem} value.

Table A2. Rate constants displaying correlations between calculated rates of ligand entry, k'_{NO}, and k'_{O2} in CerHb mutants measured at pH 7.0, 20 °C.

Mutants	k'_{NO} μM^{-1}	k'_{O2} $\mu M^{-1}s^{-1}$	$k'_{entry, CO}$ $\mu M^{-1}s^{-1}$
WT	230±30	230±17	640±40 0
Val7(B6)			
Ala7	240	250	450
Val7	110	110	140
Trp7	30	33	16
Gln44(E7)			
Phe44	110	20	26
Thr48(E11)			
Val48	80	30	27
Ala55(E18)			
Val55	140	180	260
Leu55	47	72	37
Phe55	48	61	39
Trp55	45	63	34
Leu86(G12)			
Ala86	200	230	250
Phe86	110	120	150
Trp86	68	97	30
Phe7/Phe55/Phe86	32	42	22

The numbers for wt CerHb are based on the average of all determinations for the past 6 years (≥10 separate preparations). The error in k'_{entry} is very large because F_{gem} is close to 0.0, poorly defined, and has an error of ± 60%. When F_{gem} is ≥0.1, the error in k'_{entry} diminishes greatly to ~ ± 20%.

[a] The fraction of geminate O_2 rebinding was measured using a 7 ns YAG laser light pulse.

Table 3. Bimolecular, geminate, and calculated rate parameters for CerHb mutants located in the heme pocket measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.005 (15)	590±100	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Phe10(B9)											
Ala10	70	40	1.8	6.8 (55%) 2.2 (35%) 10% slower phases	0.020	340	22	0.10	2.2	20	50
Val10	94	210	0.45	6.4	0.040 0.0070	160 910	-	-	-	-	-
Leu10	130	160	0.81	12 (81%) 2.6 (13%) 6% slower phases	0.040	300	160	0.70	110	48	15
Iso10	140	290	0.48	13	0.049 0.010	270 1,300	54	0.16	8.6	45	81
Trp10	65	8.5	7.6	12 (89%) 2.6 (11%)	0.0080	1,500	80 (84%) 13 (16%)	0.61	42	27	18

TableA3 (cont'd). Bimolecular, geminate, and calculated rate parameters for CerHb mutants located in the heme pocket measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.005 (15)	590±10 0	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Tyr11(B10)											
Ala11	120[b]	400[b]	0.30[b]	33[b]	0.015	2,200	2.8 (42%) 19 (58%)	0.28	3.4	8.8	120
Val11	150	640	0.23	40	0.021 0.00010	1,900 400,000	1.2 (23%) 14 (77%)	0.40	4.4	6.6	100
Leu11	52	610	0.085	22	0.011	2,000	24	0.66	16	8.2	33
Iso11	120	750	0.16	33	0.019 0.0010	1,700 33,000	1.3 (14%) 17 (86%)	0.44	6.5	8.3	75
Phe11	140[b]	460[b]	0.30[b]	34	0.0081	4,200	44	0.31	14	30	110
Trp11	1.3	270	0.0048	0.050	0.0080 0.0010	6.3 50	5.1 (20%) 170 (80%)	0.12	16	120	0.42

Table A3 (cont'd). Bimolecular, geminate, and calculated rate parameters for CerHb mutants located in the heme pocket measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.005 (15)	590±10 0	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Phe25(CD1)											
Val25	180	200	0.90	70	0.015	6,200 3,200	52	0.26	14	38	270
Trp25	35	99	0.35	3.2	0.17	19	200[c]	0.050	10	190	64
Phe27(CD3)											
Ala27	190	240	0.79	30	0.063	480	120	0.10	12	110	300
Val27	180	330	0.55	27	0.057	470	76	0.13	9.9	66	210
Trp27	130	83	1.6	45	0.070	640	50	0.050	2.5	48	900
Gln44(E7)											
Ala44	160	33	4.8	41	0.050 0.0040	820 10,000	81	0.090	7.3	74	460
Val44	130	84	1.5	25	0.032	780	96	0.14	14	83	180
Leu44	180	96	1.9	36	0.016	2,300	58	0.19	11	47	190
Phe44	20	170	0.12	1.3	0.054	24	80	0.050	4.0	76	26
Trp44	10 170	33 330	0.30 0.52	5.0	0.035 0.011	140 460	65	0.11	7.2	58	45
His44	85	81	1.0	4.0	0.010	400	150	0.08	12	140	50

TableA3 (cont'd). Bimolecular, geminate, and calculated rate parameters for CerHb mutants located in the heme pocket measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.005 (15)	590±100	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Lys47(E10)											
Ala47	180	350	0.51	20	0.050	400	140	0.090	13	130	220
Leu47	210	270	0.78	33	0.053	620	44	0.12	5.3	39	280
Asn47	150	270	0.56	16	0.067	240	100	0.052	5.2	95	310
Asp47	180	250	0.72	24	0.063	380	49	0.067	3.3	46	360
Phe47	190	260	0.73	26	0.065	400	25	0.065	1.6	23	400
Trp47[b]	180	180	1.0	27	0.080	340	96	0.088	8.4	88	310
Thr48(E11)											
Ala48	41[b]	0.26[b]	160[b]	2.6	0.0065	400	-	≤ 0.025	-	-	-
Val48	30[b]	0.18[b]	170[b]	3.0[b]	0.0070[b]	430[b]	67	0.11	7.4	60	27
Leu48	16	0.15	110	3.1	0.0030	1,000	70	0.11	7.7	62	28
Iso48	43	0.40	110	5.1	0.0050	1,000	110	0.11	12	98	46
Phe48	6.2	0.020	310	0.99	0.00070	1,400	46	0.73	34	12	1.4
Trp48	4.3	0.025	170	1.4	0.0035	400	13 (30%) 66 (70%)	0.72	36	14	1.9
Ala82(G8)											
Leu82	49	56	0.88	7.8	0.031 (75%) 0.0030 (25%)	250 2,600	200	0.064	13	190	120
Phe82	240	18	13	94	0.0020	47,000	74	0.32	24	50	290
Trp82	63	11	5.7	8.1	0.0030 (50%) 0.00060 (50%)	2,700 14,000	200	0.087	17	180	93

The standard deviations for the wt values for k'_{O2}, k_{O2}, and k'_{CO} were obtained from the analysis of 12 completely independent sets of experiments starting with at least 7 different expressions and purifications of CerHb. The standard deviation for k_{CO} came from the analysis of 16 different determinations. The values of K_{O2} and K_{CO} were calculated from k'_{O2}/k_{O2} and k'_{CO}/k_{CO}, and their standard deviations were calculated from the standard propagation of error formula.

[a] The fraction of geminate CO rebinding was measured using a 7 ns YAG laser light pulse.

[b] Binding parameters reported in {Pesce, 2004 #2}.

[c] Most of the amplitude change occurs during the light pulse. This k_{gem} value is an estimate.

Table A4. Bimolecular, geminate, and calculated rate parameters for CerHb multiple mutants located throughout the apolar tunnel measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.0053 (15)	590±100	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
V7A/L86A	240	190	1.3	35	0.034	1,030	8.5	0.14	1.2	7.3	250
V7A/A55F	68	87	0.78	16	0.018	890	7.3	0.48	3.5	3.8	33
V7F/A55F	63	58	1.1	15	0.013	1,200	42	0.40	17	25	38
V7F/A55W	45	37	1.2	11	0.0080	1,400	40	0.54	22	18	20
V7F/A55F/L86F	42	17	2.5	17	0.0010	17,000	39	0.79	31	8.2	22
V7W/A55F	33	9.3	3.6	12	0.0010	12,000	59	0.82	48	11	15
V7W/L86F	34	5.0	6.8	15	0.00090 0.00020	17,000 75,000	86	0.92	79	6.9	16
V7W/A55F/L86F	19	2.5	7.6	5.8	0.00026	22,000	76	0.91	69	6.8	6.4
V7W/A55W	26	10	2.6	10	0.0012	8,300	54	0.85	46	8.1	12
V7W/L86W	30	2.3	13	12	0.00030	40,000	99	0.93	92	6.9	13
V7W/A55W/L86W	23	3.0	7.6	7.2	0.00070 0.00010	10,300 72,000	59	0.93	55	4.1	7.7

The standard deviations for the wt values for k'_{O2}, k_{O2}, and k'_{CO} were obtained from the analysis of 12 completely independent sets of experiments starting with at least 7 different expressions and purifications of CerHb. The standard deviation for k_{CO} came from the analysis of 16 different determinations. The values of K_{O2} and K_{CO} were calculated from k'_{O2}/k_{O2} and k'_{CO}/k_{CO}, and their standard deviations were calculated from the standard propagation of error formula.

[a] The fraction of geminate CO rebinding was measured using a 7 ns YAG laser light pulse.

Table A5. Bimolecular, geminate, and calculated rate parameters for CerHb multiple mutants located throughout the heme pocket measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters						Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k'_{CO} $\mu M^{-1}s^{-1}$	k_{CO} s^{-1}	K_{CO} μM^{-1}	k_{gem} μs^{-1}	F_{gem} [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	230±17 (12)	190±25 (12)	1.2±0.2	32±4.7 (12)	0.054±0.0053 (15)	590±100	80±40 (8)	0.05±0.03 (8)	4.0±3.0	76±40	640±400
Y11F/Q44L	290	490	0.60	120	0.0040 0.00020	31,000 610,000	65	0.39	25	40	310
Y11F/Q44H	14	16	0.88	0.84	0.016 0.0020	53 420	54	0.16	8.6	45	5.3
Y11F/T48V	230	62	3.7	75	0.0040	19,000	65	0.33	21	44	230
Q44L/T48V	55	4.2	13	7.1	0.036 0.0069	200 1,030	95	0.16	15	80	44
Q44H/T48V	120	11	11	50, 10, 1.0, 0.20	0.010	-	120	0.36	43	77	-
Q44T/T48Q	63	360 52	0.17 1.2	4.0	0.040	100	14	0.019	0.27	14	210
Y11F/Q44L/T48V	340	270	1.6	170	0.0030	55,000	73	0.40	29	44	430
Y11F/Q44H/T48V	29	2.9	10	3.9 1.2	0.036 0.0050	-	36	0.081	2.9	33	48 15
V7W/Y11F/Q44L/T48V	82	54	1.5	40	0.00018	220,000	95	0.77	73	22	52
Y11F/Q44L/T48V/A55W	70	50	1.4	27	0.00036	75,000	50	0.85	43	7.5	32

The standard deviations for the wt values for k'_{O2}, k_{O2}, and k'_{CO} were obtained from the analysis of 12 completely independent sets of experiments starting with at least 7 different expressions and purifications of CerHb. The standard deviation for k_{CO} came from the analysis of 16 different determinations. The values of K_{O2} and K_{CO} were calculated from k'_{O2}/k_{O2} and k'_{CO}/k_{CO}, and their standard deviations were calculated from the standard propagation of error formula.

[a] The fraction of geminate CO rebinding was measured using a 7 ns YAG laser light pulse.

Table A.6 Bimolecular, geminate, and calculated rate parameters for Mb mutants located throughout the protein matrix measured at pH 7.0, 20 °C.

Mutants	Bimolecular binding parameters			Observed Geminate parameters		Calculated rate parameters		
	k'_{O2} $\mu M^{-1}s^{-1}$	k_{O2} s^{-1}	K_{O2} μM^{-1}	k_{gem} μs^{-1}	$F_{gem,\,O2}$ [a]	k_{bond} μs^{-1}	k_{escape} μs^{-1}	k'_{entry} $\mu M^{-1}s^{-1}$
WT	**16±3** [b]	**14±3**	**1.1±0.2**	**12±3**	**0.47±0.09**	**5.7±1.0**	**6.3±1.0**	**34±7**
Arg45(CD3)								
Ala45	**7.7**	**30**	**0.26**	**5.8**	**0.31**	**1.8**	**4.0**	**25**
Trp45	6.2	21	0.30	10	0.13	1.3	8.7	48
Leu61(E4)								
Ala61	**11**	**26**	**0.43**	**19**	**0.33**	**6.6**	**12**	**33**
Trp61	8.8	11	0.80	15	0.24	3.6	11	37
Leu69(E12)								
Ala69	28	13	2.2	22	0.46	10	12	61
Trp69	13	5.2	2.5	32	0.61	20	13	21
Ile75(E18)								
Ala75	22	8.5	2.6	16	0.35	5.6	10	63
Trp75	16	8.8	1.8	16	0.43	6.9	9.1	37
Leu135(H12)								
Val135	21	8.1	2.6	20	0.40	8.0	12	53
Trp135	23	6.5	3.5	18	0.33	5.9	12	70

The value of K_{O2} were calculated from k'_{O2}/k_{O2}, and its standard deviation was calculated from the standard propagation of error formula. When fitting the O_2 geminate data, single and double exponential fits were examined. In the case of k_{gem} values, estimates from 1 exponential fits were used as final values. In the case of F_{gem} values, 2 exponential fits were used to estimate total $F_{gem,O2}$.

[a] The fraction of geminate O_2 rebinding was measured using a 7 ns YAG laser light pulse.

[b] Rate constants in bold are taken from {Scott, 2001 #14}.